National Institute for Health and Care Excellence

Version 1.0

Depression in adults: treatment and management

Full guideline

NICE Guideline <…>

Methods, evidence and recommendations

18 July 2017

Draft for Consultation

Developed by the National Guideline Alliance, hosted by the Royal College of Obstetricians and Gynaecologists

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Contents

Update 2017

Update 2017

Update 2017

Update 2017

Update 2017

Update 2017

Update 2017

Update 2017

Update 2017

Update 2017

Update 2017

Update 2017

Guideline Committee members and review team

Guideline Committee 2017

Navneet Kapur (Chair, Guideline Committee)
Professor of Psychiatry and Population Health, University of Manchester and Honorary Consultant in Psychiatry, Greater Manchester Mental Health and Social Care Trust

June Brown
Senior Lecturer in Clinical Psychology, Institute of Psychiatry, King's College London

Carolyn Chew Graham
Professor of General Practice Research, Keele University; Honorary Professor of Primary Care, University of Manchester; Honorary Professor of Primary Care Mental Health, South Staffs & Shropshire Foundation Trust (SSSFT); Principal in General Practice, NHS Manchester; Honorary Consultant in Primary Care, NHS Manchester

Jeremy Clarke
Psychological therapist, Newham Psychological Therapies, ELFT; Senior Accredited DIT Supervisor, Southwark IAPT, Wandsworth IAPT, Sutton and Merton IAPT; Consultant Mental Health Advisor, Greater Manchester Combined Authority); Research Associate, Centre for Humanities Engaging Science and Society, University of Durham

Sinead Dervin
Senior Commissioning Manager, Mental Health, NHS England

David Ekers
Nurse Consultant/Senior Visiting Research Fellow, Tees Esk and Wear Valleys NHS Trust/University of York (Mental Health and Addictions Research Group)

Simon Gilbody
Professor of Psychological Medicine and Health Services Research, University of York

David Hewison
Head of Research; Consultant to 'Couple Therapy for Depression' Training, Tavistock Relationships; Programme Leader & Supervisor, Professional Doctorate in Couple Psychotherapy, TCCR/UEL

Tony Kendrick
Professor of Primary Care, University of Southampton

Louise O'Connor
Representing Service User and Carer views

Neil Nixon
Consultant Psychiatrist and Director of Medical Education, Nottinghamshire Healthcare NHS Trust; Honorary Associate Professor, University of Nottingham

Catherine Ruane
Representing Service User and Carer views

Prishah Shah
Representing Service User and Carer views

Jennifer Speller
Head of Service Transformation (acting), Islington Clinical Commissioning Group

Update 2017

David Taylor
Director of Pharmacy and Pathology, South London and Maudsley NHS Foundation Trust; Professor of Pharmacology, King's College London

Edward Watkins
Professor of Experimental and Applied Clinical Psychology, University of Exeter; Chartered Clinical Psychologist; Director of Sir Henry Wellcome Building for Mood Disorders Research; Director of Research for DClinPsy Programme; Director of SMART Lab

Philip Wilkinson
Consultant in Psychiatry of Old Age, Oxford Health NHS Foundation Trust and Honorary Senior Clinical Lecturer, University of Oxford

Review team 2017

Lauren Becker
Research Assistant, National Collaborating Centre for Mental Health (NCCMH) (to March 2016)

Angela Bennett
Guideline Lead, National Guideline Alliance (NGA) (from July 2016)

Bishal Bhandari
Assistant Systematic Reviewer, NGA (October 2016 to January 2017)

Gemma Halliday
Research Assistant, NCCMH (to March 2016)

Sally Humphreys
Project Manager, NGA (from April 2016)

Ifigeneia Mavranezouli
Senior Health Economist, NGA

Odette Megnin Viggars
Senior Systematic Reviewer, NGA

Maryla Moulin
Senior Project Manager, NCCMH, (until March 2016)

Stephen Pilling
Clinical Adviser, NGA

Emma Seymour
Systematic Reviewer, NGA (from July 2016)

Sarah Stockton
Senior Information Scientist, NGA

Iona Symington
Systematic Reviewer, NGA (until July 2016)

Update 2017

Guideline Development Group 2009

Ian Anderson (Chair, Guideline Development Group)
Professor of Psychiatry, University of Manchester

Stephen Pilling
Director, NCCMH; Director, Centre for Outcomes Research and Effectiveness, University College London

Alison Barnes
Service User Member

Linda Bayliss
Research Assistant (May 2008 to August 2008), NCCMH

Victoria Bird
Research Assistant, NCCMH

Rachel Burbeck
Lead Systematic Reviewer, NCCMH

Carolyn Chew-Graham
General Practitioner and Senior Lecturer in Primary Care, University of Manchester

Jeremy Clarke
Psychological Therapist, Lambeth Primary Care Trust

Matthew Dyer
Health Economist, NCCMH

Esther Flanagan
Project Manager (2009), NCCMH

Catherine Harris
Carer member and Local Councillor

Sarah Hopkins
Project Manager (until 2008), NCCMH

Mark Kenwright
Consultant Cognitive Behavioural Psychotherapist, Ealing Cognitive Behavioural Therapy Service

Willem Kuyken
Professor of Clinical Psychology and Co-Director, Mood Disorders Centre, School of Psychology, University of Exeter

Angela Lewis
Research Assistant, NCCMH

Glyn Lewis
Professor of Psychiatric Epidemiology, University of Bristol

Ryan Li
Project Manager (2008), NCCMH

Brendan Masterson
Clinical Nurse Leader, Affective Disorders Unit, Bethlem Royal Hospital

Nick Meader
Systematic Reviewer, NCCMH

Alan Meudell
Service User Member, Healthy Minds at Work

Alex Mitchell
Consultant Psychiatrist and Honorary Lecturer in Liaison Psychiatry, University of Leicester

Richard Moore
Clinical Psychologist, Cambridge and Peterborough NHS Foundation Trust

Suffiya Omarjee
Health Economist, NCCMH

Carol Paton
Chief Pharmacist, Oxleas NHS Foundation Trust

Alejandra Perez
Systematic Reviewer, NCCMH

Peny Retsa
Health Economist (until 2008), NCCMH

Maria Rizzo
Research Assistant, NCCMH

Jennie Robertson
Research Assistant (from September 2008), NCCMH

Rob Saunders
Research Assistant (2008), NCCMH

Christine Sealey
Centre Manager, NCCMH

Beth Shackleton
Project Manager (until 2008), NCCMH

Thomas Shackleton
General Practitioner, Suffolk

Sarah Stockton
Senior Information Scientist, NCCMH

Clare Taylor
Editor, NCCMH

Jane Wood
Nurse, Strategic Development Manager, Mental Health, Leeds Primary Care Trust

Guideline Development Group 2004

David Goldberg (Chair, Guideline Development Group)
Emeritus Professor of Psychiatry, Institute of Psychiatry, King's College, London

Stephen Pilling (Facilitator, Guideline Development Group)
Co-director, NCCMH; Director, Centre for Outcomes, Research and Effectiveness;
Consultant Clinical Psychologist, Camden and Islington Mental Health and Social Care Trust, London

Tim Kendall (Facilitator, Guideline Development Group)
Co-director, NCCMH; Deputy Director, Royal College of Psychiatrists' Research Unit;
Medical Director and Consultant Psychiatrist, Community Health Sheffield NHS Trust

Nicol Ferrier (Lead, Topic Group on Pharmacology)
Head of School of Neurology, Neurobiology and Psychiatry, University of Newcastle

Ted Foster
Patient Representative National Advisory Panel Member – Mind Link

John Gates
Patient Trustee, National Mind; Chair, Redcar and Cleveland Mind

Paul Gilbert (Lead, Topic Group on Psychology)
Mental Health Research Unit, Kingsway Hospital, University of Derby

Paul Harvey
General Practitioner, Devonshire Green Medical Centre, Sheffield

Ian Hughes
Consultant Clinical Psychologist, Cardiff and Vale NHS Trust

Carol Paton
Chief Pharmacist, Oxleas NHS Trust, south east London

Simon Rippon
Programme Co-ordinator, National Institute for Mental Health in England, Northwest Development Centre

Kay Sheldon
Patient Representative

Douglas Turkington
Senior Lecturer in Liaison Psychiatry, University of Newcastle-Upon-Tyne, Royal Victoria Infirmary; Consultant Psychiatrist, Newcastle, North Tyneside and Northumberland Mental Health NHS Trust

André Tylee (Lead, Topic Group on Service Interventions)
Professor of Primary Care Mental Health, Institute of Psychiatry, London

Rachel Burbeck
Lead Systematic Reviewer, NCCMH

Michelle Clark
Project Manager (to September 2003), NCCMH

Cesar de Olivera
Systematic Reviewer, NCCMH

Catherine Pettinari
Senior Project Manager (from September 2003), NCCMH

Preethi Premkumar
Research Assistant, NCCMH

Judit Simon
Health Economist, NCCMH

Clare Taylor
Editor, NCCMH

Lisa Underwood
Research Assistant, NCCMH

Craig Whittington
Senior Systematic Reviewer, NCCMH

1 **Heather Wilder**
2 Information Scientist, NCCMH

Acknowledgements

Acknowledgements 2017

The Guideline Committee (GC) and the NGA would like to thank the following people:

Adam Chekroud, Department of Psychology, Yale University, New Haven, CT, USA

NICE Guidelines Technical Support Unit (TSU)

Tony Ades, Senior Adviser, TSU, University of Bristol

Sofia Dias, Director, TSU, University of Bristol

Edna Keeney, Scientific-coordinator, TSU, University of Bristol

Nicky Welton, Deputy Director, TSU, University of Bristol

Centre for Outcomes Research and Effectiveness, University College London

Katherine Clarke, Research Assistant, Centre for Outcomes Research and Effectiveness, University College London

Suzan Hassan, Research Assistant, Centre for Outcomes Research and Effectiveness, University College London

Bianca Reveruzzi, Centre for Outcomes Research and Effectiveness, University College London

Rob Saunters, Research Associate, Centre for Outcomes Research and Effectiveness, University College London

Cochrane

Hanna Bergman, Systematic Reviewer, Cochrane Response

Rachel Churchill, Coordinating Editor, Cochrane Common Mental Disorders Group

Christine Clar, Systematic Reviewer, Cochrane Response

Sarah Dawson, Trials Search Co-ordinator, Cochrane Common Mental Disorders Group

Rachel Marshall, Executive Editor, Cochrane Response

Loukia Spineli, Systematic Reviewer, Cochrane Response

Rachel Wheeler, Freelance Consultant, Cochrane Response

Charlotte Pestridge, CEO Cochrane Innovations

Karla Soares-Weiser, Deputy Editor-in-Chief, Cochrane and Cochrane Innovations

Update 2017

Acknowledgements 2009

Editorial assistance

Nuala Ernest, Assistant Editor, NCCMH

Marie Halton, Research Assistant, NCCMH

Acknowledgements 2004

Guideline Review Panel

The Guideline Review Panel is an independent panel that oversees the development of the guideline and takes responsibility for monitoring its quality. The Panel includes experts on guideline methodology, health professionals and people with experience of the issues affecting patients and carers. The members of the Guideline Review Panel were as follows.

Chaand Nagpaul, GP, Stanmore

John Seddon, Patient Representative

Kenneth Wilson , Professor of Psychiatry of Old Age and Honorary Consultant Psychiatrist, Cheshire and Wirral Partnership NHS Trust

Paul Rowlands, Consultant Psychiatrist, Derbyshire Mental Health Services Mental Health Care Trust

Roger Paxton, R&D Director, Newcastle, North Tyneside and Northumberland Mental Health NHS Trust

1 Preface

This guidance is a partial update of NICE clinical guideline CG90 (NICE 2009) and will replace it..

This guideline was first published in December 2004 (NICE 2004, NICE 2004) and updated in 2009 (NICE 2009, NICE 2010). The previous guidelines and this update have been developed to advise on the treatment and management of depression. The guideline recommendations in the update have been developed by a multidisciplinary team of healthcare professionals, service users, carers and 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 people with depression while also emphasising the importance of the experience of care for them and their carers.

The present guideline updates most areas of the previous guideline. It should be noted that because the NICE guideline on service user experience in adult mental health services (NICE 2011, NICE 2012) covers the experience of care for people accessing mental health services (including people with depression), Chapter 4 on Experience of care was not updated from 2009, nor was the section on identification. The superseded text from the 2009 guideline can be seen in Appendix V. The 2009 guideline was divided into chapters on types of intervention, whereas the 2017 guideline has chapters on the treatment and management of different aspects of the condition.

New and updated recommendations have been included on organisation and delivery of services, access to services, the treatment of new depressive episodes, further-line treatment of depression, chronic depression, depression with co-morbidities and relapse prevention. Recommendations in the previous guideline were reviewed for their current relevance and terminology. See Appendix A for more details on the scope of this update.

Update 2017

Recommendations are marked to indicate the year of the last evidence review:

- [2009] or [2004] if the evidence has not been reviewed since the original guideline.
- [2009 or 2004, amended 2017] if the evidence has not been reviewed, but an essential change has been made that affects the meaning of the recommendation.
- [2017] if the evidence has been reviewed but no change has been made to the recommendation.
- [new 2017] if the evidence has been reviewed and the recommendation has been updated or added.

Where recommendations are shaded in grey and end [2004] or [2009] the evidence has not been updated since the original guideline. Yellow shading in these recommendations indicates where wording changes have been made for the purposes of clarification only.

You are invited to comment on the new and updated recommendations in this guideline only. These are marked as [2017] if the evidence has been reviewed but no change has been made to the recommendation or [new 2017] if the evidence has been reviewed and the recommendation has been added or updated.

Appendix V3 contains recommendations from the 2009 guideline that NICE proposes deleting in the 2017 update. This is because the evidence has been reviewed and the recommendation has been updated or because NICE has updated other relevant guidance and has replaced the original recommendations. Where there are replacement recommendations, details are provided. Where there is no replacement recommendation, an explanation for the proposed deletion is given. You are invited to comment on the deleted recommendations as part of the consultation on the 2017 update.

The original NICE guideline and supporting documents are available from:
www.nice.org.uk/guidance/CG90

Although the evidence base is rapidly expanding there are a number of major gaps, and further revisions of this guideline will incorporate new scientific evidence as it develops. The guideline makes a number of research recommendations specifically to address gaps in the evidence base. In the meantime, it is hoped that the guideline will assist clinicians, people with depression and their carers by identifying the merits of particular treatment approaches where the evidence from research and clinical experience exists.

1.1 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 Committee (GC).

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.

Update 2017

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 methodology used in the development of the guideline, the generalisability of research findings and the uniqueness of individuals with depression.

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 (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 people with depression. 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 the circumstances of the individual, in consultation with the person with depression or their 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 clinical 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?

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, 3 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 clinical guidelines focused upon the overall treatment and management of a specific condition.

Update 2017

1.1.4 From clinical guidelines to local implementation

Once a clinical 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, people with depression and their 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 1999) 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 Depression in Adults guideline

1.2.1 Who has developed this guideline?

This guideline has been commissioned by NICE and was initially developed within the National Collaborating Centre for Mental Health (NCCMH). The NCCMH was 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 was funded by NICE and 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.

On 1 April 2016 the NCCMH was amalgamated into the National Guideline Alliance (NGA) at the Royal College of Obstetricians and Gynaecologists, along with the National Collaborating Centre for Women and Children's Health and the National Collaborating Centre for Cancer.

The technical team provided leadership and support throughout the process of guideline development, undertaking systematic searches, information retrieval, appraisal, systematic reviewing of the evidence and training for the GC in the process of guideline development. Service users and carers received additional training and support from the NICE Public Involvement Programme and the NICE Guidelines Technical Advisor provided methodological advice and assistance.

All GC members made formal declarations of interest at the outset, which were updated at every GC meeting. The GC met a total of 14 times throughout the process of guideline development. The GC was supported at all stages by the technical team, with additional expert advice from special advisers where needed. The committee oversaw the synthesis of research evidence and all statements and recommendations in this guideline have been generated and agreed by the whole GC.

Update 2017

1.2.2 For whom is this guideline intended?

This guideline is relevant for adults with depression as the primary diagnosis 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, adults with depression.

The guideline will also be relevant to the work, but will not cover the practice, of those in:

- occupational health services
- social services
- forensic services
- the independent sector.

The experience of depression can affect the whole family and often the community. The guideline recognises the role of both in the treatment and support of people with depression.

1.2.3 Specific aims of this guideline

The guideline makes recommendations for the treatment and management of depression. It aims to:

- improve access and engagement with treatment and services for people with depression
- evaluate the role of specific psychological and psychosocial interventions in the treatment of depression
- evaluate the role of specific pharmacological interventions in the treatment of depression
- evaluate the role of specific service-level interventions for people with depression

- integrate the above to provide best-practice advice on the care of people with depression and their family and carers
- 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 3 chapters provide an introduction to guidelines, the topic of depression and the methods used to update this guideline. The following chapters provide the evidence that underpins the recommendations about the treatment and management of depression.

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 the review protocol and studies included in the review. Clinical evidence summaries are used to summarise the data presented. Health economic evidence is then presented (where appropriate), followed by the recommendations related to each topic and a section (from evidence to recommendations) that draws together the clinical and health economic evidence and provides a rationale for the recommendations. In the appendices, further details are provided about included/excluded studies, the evidence, and the previous guideline methodology (see Table 1 for details). Where meta-analyses were conducted, the data are presented using forest plots.

Table 1: Appendices

Content	Appendix
Scope for the development of the clinical guideline	Appendix A
Declarations of interests by Guideline Committee members	Appendix B
Special advisers to the Guideline Committee	Appendix C
Stakeholders	Appendix D
Researchers contacted to request information about unpublished or soon-to-be published studies	Appendix E
Review questions and review protocols	Appendix F
Research recommendations	Appendix G
Search strategies – clinical evidence	Appendix H
Search strategies – economic evidence	Appendix I
Study characteristics	Appendix J
Clinical evidence – flow charts	Appendix K
Clinical evidence – GRADE evidence profiles	Appendix L
Clinical evidence – forest plots	Appendix M
Clinical evidence – network meta-analysis: bias adjustment methods and results	Appendix N
Economic evidence – flow chart	Appendix O
Economic evidence – health economic checklists	Appendix P
Economic evidence – evidence tables	Appendix Q
Economic evidence – economic profiles	Appendix R
Economic evidence – list of excluded studies	Appendix S
Studies and data included in the network meta-analyses	Appendix T
Study references from 2004 and 2009 guidelines	Appendix U

Update 2017

Content	Appendix
Superseded text from CG90 guideline	Appendix V
Network meta-analysis inconsistency checks - direct vs NMA results	Appendix W

1

Update 2017

2 Introduction

This guideline is concerned with the treatment and management of adults with a primary diagnosis of depression in primary and secondary care. The terminology and diagnostic criteria used for this heterogeneous group of related disorders have changed over the years, and the 2004 guideline related only to those identified by The ICD–10 Classification of Mental and Behavioural Disorders (ICD–10) WHO (1992) as having a depressive episode (F32 in the ICD–10), recurrent depressive episode (F33) or mixed anxiety and depressive disorder (F41.2). In the 2009 guideline update the scope was widened to cover the substantial proportion of people who present with less severe forms of depression. Therefore, this updated guideline covers 'subthreshold depressive symptoms', which fall below the criteria for major depression (and which do not have a coding in ICD–10), and subthreshold depressive symptoms persisting for at least 2 years (dysthymia; F34.1).

It should, however, be noted that much of the research forming the evidence base from which this guideline is drawn has used a different classificatory system – the Diagnostic and Statistical Manual of Mental Disorders of the American Psychiatric Association, currently in its fifth edition (DSM–5) (American Psychiatric Association (2013). The two classificatory systems, while similar, are not identical especially with regard to definitions of severity. After considerable discussion the GC took the decision to base the guidelines on the DSM–IV-TR (see Section 2.1.5). This covers major depressive disorder single episode (296.2) and recurrent (296.3) together with dysthymic disorder (300.4), and contains research criteria for minor depressive disorder (APA 2000c). The effect of this change in practice is discussed in Section 2.1.5. The core criterion symptoms applied to the diagnosis of major depressive episode, and the requisite duration of at least 2 weeks, have not changed from DSM-IV to DSM-V. The requirement for clinically significant distress or impairment in social, occupational, or other important areas of life is also unchanged, although this is now listed as Criterion B rather than Criterion C. In DSM-IV, there was an exclusion criterion for a major depressive episode that was applied to depressive symptoms lasting less than 2 months following the death of a loved one, but this exclusion is omitted in DSM-5 (APA 2014). DSM-5 also reclassified what was called dysthymia in DSM-IV as persistent depressive disorder, which includes both chronic major depressive disorder and the previous dysthymic disorder (APA 2014).

The guideline does not address the management of depression in children and adolescents, depression in bipolar disorder, depression occurring in both antenatal and postnatal periods, or depression associated with chronic physical health problems, all of which are covered by separate guidelines:

- depression in children and young people: identification and management; NICE (2005)
- bipolar disorder: assessment and management; NICE (2014)
- antenatal and postnatal mental health: clinical management and service guidance; NICE (2014)
- depression in adults with a chronic physical health problem: recognition and management; NICE (2010).

The guideline update does cover psychotic symptoms occurring within the context of an episode of depression (depression with psychotic symptoms), but not depression occurring in a primary psychotic illness, such as schizophrenia or dementia.

Update 2017

2.1 What is depression?

2.1.1 Symptoms, presentation and pattern of illness

Depression refers to a wide range of mental health problems characterised by the absence of a positive affect (a loss of interest and enjoyment in ordinary things and experiences), low mood and a range of associated emotional, cognitive, physical and behavioural symptoms. Distinguishing the mood changes between clinically significant degrees of depression (for example, major depression) and those occurring 'normally' remains problematic and it is best to consider the symptoms of depression as occurring on a continuum of severity (Lewinsohn et al. 2000). The identification of major depression is based not only on its severity but also on persistence, the presence of other symptoms, and the degree of functional and social impairment. However, there appears to be no hard-and-fast 'cut-off' between 'clinically significant' and 'normal' degrees of depression; the greater the severity of depression, the greater the morbidity and adverse consequences (Lewinsohn et al. 2000, Kessing 2007). When taken together with other aspects that need to be considered, such as duration, stage of illness and treatment history, there are considerable problems when attempting to classify depression into categories (see Section 2.1.5).

Commonly, mood and affect in a major depressive illness are unreactive to circumstance, remaining low throughout the course of each day, although for some people mood varies diurnally, with gradual improvement throughout the day only to return to a low mood on waking. For others, a person's mood may be reactive to positive experiences and events, although these elevations in mood are not sustained, with depressive feelings re-emerging, often quickly (Andrews and Jenkins 1999).

Behavioural and physical symptoms typically include tearfulness, irritability, social withdrawal, an exacerbation of pre-existing pains, pains secondary to increased muscle tension (Gerber et al. 1992), a lack of libido, fatigue and diminished activity, although agitation is common and marked anxiety frequent. Typically there is reduced sleep and lowered appetite (sometimes leading to significant weight loss), but for some people it is recognised that sleep and appetite are increased. A loss of interest and enjoyment in everyday life, and feelings of guilt, worthlessness and that one deserves punishment, are common, as are lowered self-esteem, loss of confidence, feelings of helplessness, suicidal ideation and attempts at self-harm or suicide. Cognitive changes include poor concentration and reduced attention, pessimistic and recurrently negative thoughts about oneself, one's past and the future, mental slowing and rumination (Cassano and Fava 2002).

Depression is often accompanied by anxiety, and in these circumstances one of three diagnoses can be made: (1) depression; (2) anxiety; or (3) mixed depression and anxiety when both are below the threshold for either disorder, dependent upon which constellation of symptoms dominates the clinical picture. In addition, the presentation of depression can vary with age with the young showing more behavioural symptoms and older adults more somatic symptoms and fewer complaints of low mood (Serby and Yu 2003).

Major depression is generally diagnosed when a persistent low mood and an absence of positive affect are accompanied by a range of symptoms, the number and combination needed to make a diagnosis being operationally defined (ICD–10, WHO 1992; DSM–IV, APA 1994).

Some people are recognised as showing an atypical presentation with reactive mood, increased appetite, weight gain and excessive sleepiness together with the personality feature of sensitivity to rejection (Quitkin et al. 1991) and this is classified as major depression with atypical features in DSM–IV (APA 1994). The definition of atypical depression has changed over time and it is not specifically recognised in ICD–10.

Update 2017

Some patients have a more severe and typical presentation, including marked physical slowness (or marked agitation), complete lack of reactivity of mood to positive events, and a range of somatic symptoms, including appetite and weight loss, reduced sleep with a particular pattern of waking early in the morning and being unable to get back to sleep. A pattern of the depression being substantially worse in the morning (diurnal variation) is also commonly seen. This presentation is referred to as major depression with melancholic features in DSM–IV and a depressive episode with somatic symptoms in ICD–10.

People with severe depression may also develop psychotic symptoms (hallucinations and/or delusions), most commonly thematically consistent with the negative, self-blaming cognitions and low mood typically encountered in major depression, although others may develop psychotic symptoms unrelated to mood (Andrews and Jenkins 1999). In the latter case, these mood-incongruent psychotic symptoms can be hard to distinguish from those that occur in other psychoses such as schizophrenia.

2.1.2 Course and prognosis

The average age of the first episode of major depression occurs in the mid-20s and, although the first episode may occur at any time from early childhood through to old age, a substantial proportion of people have their first depression in childhood or adolescence (Fava and Kendler 2000). Just as the initial presentation and form of a depressive illness varies considerably, so too does the prodromal period. Some individuals experience a range of symptoms in the months prior to the full illness, including anxiety, phobias, milder depressive symptoms and panic attacks; others may develop a severe major depressive illness fairly rapidly, not uncommonly following a major stressful life event. Sometimes somatic symptoms dominate the clinical picture leading the clinician to investigate possible underlying physical illness until mood changes become more obvious.

Update 2017

Although depression has been thought of as a time-limited disorder, lasting on average 4 to 6 months with complete recovery afterwards, it is now clear that incomplete recovery and relapse are common. The WHO study of mental disorders in 14 centres across the world found that 50% of patients still had a diagnosis of depression 1 year later (Simon et al. 2002) and at least 10% had persistent or chronic depression (Kessler et al. 2003). At least 50% of people, following their first episode of major depression, will go on to have at least one more episode (Kupfer 1991) and, after the second and third episodes, the risk of further relapse rises to 70 and 90%, respectively (Kupfer 1991). People with early onset depression (at or before 20 years of age) and depression occurring in old age have a significantly increased vulnerability to relapse (Giles et al. 1989, Mitchell and Subramaniam 2005). Thus, while the outlook for a first episode is good, the outlook for recurrent episodes over the long term can be poor with many patients experiencing symptoms of depression over many years (Akiskal 1986).

Sometimes, recurrent episodes of depression will follow a seasonal pattern which has been called ‘seasonal affective disorder’ (SAD; Rosenthal et al. 1984). DSM–IV includes criteria for a seasonal pattern whereas only provisional criteria are given in the research version of ICD–10. Although a seasonal pattern can apply to both recurrent depression and bipolar disorder it appears most common in the former (70 to 80%, Rodin and Thompson 1997, Westrin and Lam 2007), with recurrent winter depression far more common than recurrent summer episodes (Rodin and Thompson 1997, Magnusson and Partonen 2005).

Depression with a seasonal pattern refers to depression that occurs repeatedly at the same time of year (not accounted for by psychosocial stress) with remission in between and without a lifetime predominance of non-seasonal depression. Decreased activity is reported as nearly always present and atypical depressive symptoms, particularly increased sleep, weight gain and carbohydrate craving are common (Magnusson and Partonen 2005). The onset is reported as usually in the third decade and is more common in the young (Rodin and Thompson 1997, Magnusson and Partonen 2005). Surveys in the UK have found a

surprisingly high prevalence in general practitioner (GP) practice attendees ranging from 3.5% in Aberdeen (Eagles et al. 1999) to 5.6% in southern England (Thompson et al. 2004). However, the validity of 'seasonal affective disorder' has been poorly accepted in Europe and may be an extreme form of a dimensional 'seasonality trait' rather than a specific diagnosis (Kasper et al., 1989). Some patients with non-seasonal mood disorders also report seasonal variation (Bauer and Dunner 1993) and this also occurs in other disorders such as anxiety and eating disorders (Bauer and Dunner 1993, Magnusson and Partonen 2005). After 5 to 11 years' follow-up, approximately half of those with continuing depressive episodes no longer display a seasonal pattern (Magnusson and Partonen 2005). A recent cross-sectional survey of 1754 US adults found depression on the PHQ-8 questionnaire to be unrelated to latitude, season, or sunlight (Traffanstedt et al. 2016).

Up to 10% of people with depression subsequently experience hypomanic/manic episodes (Kovacs 1996), which emphasises the need to question patients about a history of elevated mood and to be alert to new episodes occurring.

In a large WHO naturalistic study in 15 cities around the world, episodes of depression that were either untreated by the GP or missed entirely had the same outlook as treated episodes of depression; however, they were milder at index consultation (Goldberg et al. 1998). Thompson et al. (2001) also found that unrecognised cases were relatively mild, and GPs were better at recognising moderate to severe depression. A small longitudinal study (Kessler et al. 2002) found that the majority of undetected people either recovered or were diagnosed during the follow-up period; nevertheless, nearly 20% of the identified cases in this study remained undetected and unwell after 3 years.

2.1.3 Disability and mortality

Update 2017

Depression is the most common mental disorder in community settings and is a major cause of disability across the world. In 1990 it was the fourth most common cause of loss of disability-adjusted life years (DALYs) in the world, and it is projected to become the second most common cause by 2020 (World Bank 1993). In 1994, it was estimated that about 1.5 million DALYs were lost each year in the West as a result of depression (Murray et al. 1994). It is even more common in the developing world (for a review, see Institute of Medicine 2001). There is a clear dose–response relationship between illness severity and the extent of disability (Ormel and Costa e Silva 1995) and onsets of depression are associated with onsets of disability, with an approximate doubling of both social and occupational disability (Ormel et al. 1999). Apart from the subjective experiences of people with depression, the impact on social and occupational functioning, physical health and mortality is substantial. Depressive illness causes a greater decrement in health state than the major chronic physical illnesses: angina, arthritis, asthma and diabetes (Moussavi et al. 2007). Emotional, motivational and cognitive effects substantially reduce a person's ability to work effectively, with losses in personal and family income as well as lost contribution to society in tax revenues and employment skills. The King's Fund estimated that in the UK 1.45 million people would have depression by 2026, and the total cost to the nation would exceed GBP 12 billion per year, including prescriptions, inpatient and outpatient care, supported accommodation, social services and lost employment (McCrone 2008). Wider social effects include: greater dependence upon welfare and benefits, with loss of self-esteem and self-confidence; social impairments, including reduced ability to communicate and sustain relation- ships during the illness with knock-on effects after an episode; and longer-term impairment in social functioning, especially for those who have chronic or recurrent disorders. The stigma associated with mental health problems generally (Sartorius 2002), and the public view that others might view a person with depression as unbalanced, neurotic and irritating (Priest et al. 1996), may partly account for the reluctance of people with depression to seek help (Bridges and Goldberg 1987).

Depression can also exacerbate the pain, distress and disability associated with physical health problems as well as adversely affecting outcomes. Depression combined with chronic

physical health problems incrementally worsens health compared with physical disease alone or even combinations of physical diseases (Moussavi et al. 2007). In addition, for a range of physical health problems, findings suggest an increased risk of death when comorbid depression is present (Cassano and Fava 2002). In coronary heart disease, for example, depressive disorders are associated with an 80% increased risk, both of its development and of subsequent mortality in established disease, at least partly through common contributory factors (Nicholson et al. 2006). There is another guideline on depression in adults with a chronic physical health problem to accompany this guideline (NCCMH 2010, NICE 2009).

Suicide accounts for nearly 1% of all deaths and nearly two-thirds of this figure occur in people with depression (Sartorius 2001). Looked at another way, having depression leads to over a four-times higher risk of suicide compared with the general population, which rises to nearly 20 times in the most severely ill (Bostwick and Pankratz 2000). Sometimes depression may also lead to acts of violence against others and may even include homicide. Marital and family relationships are frequently negatively affected, and parental depression may lead to neglect of children and significant disturbances in children (Ramachandani and Stein 2003).

2.1.47 Incidence and prevalence

Worldwide estimates of the proportion of people who are likely to experience depression in their lifetime vary widely between studies and settings, but the best estimates lie between about 4 and 10% for major depression, and between about 2.5 and 5% for dysthymia (low grade chronic depressive symptoms) (Waraich et al. 2004) with disparities attributable to real differences between countries and the method of assessment. The estimated point one-week prevalence for a depressive episode (F32/33, ICD–10; WHO 1992) among 16- to 74-year-olds in the UK in 2014 was 3.3%, but, if the broader and less specific category of 'common mental disorders not otherwise specified' (representing mixed depression and anxiety) (F41.2, ICD–10, WHO 1992) was included, this figure rose dramatically to 11.1% (McManus et al. 2016).

Update 2017

Prevalence has consistently been found to be between 1.5 and 2.5 times higher in women than men and has also been fairly stable in the age range of 18 to 64 years (Waraich et al. 2004), although in the most recent UK survey cited above female preponderance was only marked for a depressive episode in those under 35 years whereas for mixed anxiety and depression it was across the age range. Compared with adults without a neurotic disorder, those with a depressive episode or mixed anxiety and depression were more likely to be aged between 35 and 54 years, separated or divorced and living alone or as a lone parent. This pattern was broadly similar between men and women (Singleton et al. 2001).

A number of socioeconomic factors significantly affected prevalence in the UK survey: those with a depressive episode were more likely than those without 'neurotic disorders' (depressive or anxiety disorders) to be unemployed, to belong to social classes 4 and below, to have lower predicted intellectual function, to have no formal educational qualifications and to live in local authority or Housing Association accommodation, to have moved three or more times in the last 2 years and to live in an urban environment (Singleton et al. 2001).

No significant effect of ethnic status on prevalence of a depressive episode or mixed anxiety and depression was found, although numerically there was a higher proportion of South Asians in those with depressive or anxiety disorders than in those without (Singleton et al. 2001). Migration has been high in Europe in the last 2 decades, but data on mental health is scarce and results vary between migrant groups (Lindert et al. 2008).

An illustration of the social origins of depression can be found in a general practice survey in which 7.2% (range 2.4 to 13.7%, depending upon the practice) of consecutive attendees had a depressive disorder. Neighbourhood social deprivation accounted for 48.3% of the variance among practices and the variables that accounted for most of that variance were:

the proportion of the population having no or only one car; and neighbourhood unemployment (Ostler et al. 2001).

There is concern that depression might be increasing in prevalence worldwide, although the evidence is mixed. Epidemiological surveys suggest prevalence increased from the early 1990s up until 2004, at least in the USA (Hasin et al. 2005, Kessler et al. 2005, Eaton et al. 2007). Overall rates in the UK did not appear to have risen at least up until 2007 (Singleton et al. 2003, McManus et al. 2009), although there was limited evidence of an increase among women (Spiers et al. 2012). Major depressive disorder (MDD) moved up from 15th to 11th in the global ranking of disorders by disability adjusted life years between 1990 and 2010 (a 37% increase) (Murray et al. 2012), but this change in ranking was actually due to population growth and ageing – prevalence of MDD was found to have decreased slightly over the 20 year period (Ferrari et al. 2013).

Kendrick et al. (2015) found that the economic recession of 2008 was followed by a modest increase in the incidence and prevalence of recorded depression in English general practices over the next five years, more in men than women, more in deprived areas, and associated with a rise in unemployment. A rise in the annual incidence of first-ever depression from 0.9% to 1% was seen in younger adults, and the overall annual prevalence rose slightly from 3.8% to 3.95% (Kendrick et al., 2015). This finding was consistent with previous findings for suicide (Barr et al. 2012, Coope et al. 2014). Youth unemployment, particularly in men, was a feature of the 2008 economic recession (Bell and Blanchflower 2011), and associations were found by Barr et al. (2012) between regional unemployment and suicide rates, while Coope et al. (2014) found increased suicide rates among men aged 35–44 years mirrored recession-related unemployment.

The evidence therefore overwhelmingly supports the view that the prevalence of depression, however it is defined, varies according to gender, and social and economic factors.

Update 2017

2.1.5 Diagnosis

In recent years there has been a greater recognition of the need to consider depression that is ‘subthreshold’; that is, where the depression does not meet the full criteria for a depressive/major depressive episode. Subthreshold depressive symptoms cause considerable morbidity and human and economic costs, and are more common in those with a history of major depression as well as being a risk factor for future major depression (Rowe and Rapaport 2006).

There is no accepted classification for subthreshold depression in the current diagnostic systems, with the closest being minor depression (a research diagnosis in DSM–IV). At least two but less than five symptoms are required and it overlaps with ICD–10 mild depressive episode with four symptoms. Given the practical difficulty and inherent uncertainty in deciding thresholds for significant symptom severity and disability, there is no natural discontinuity between subthreshold depressive symptoms and ‘mild major’ depression in routine clinical practice.

Diagnostic criteria and methods of classification of depressive disorders have changed substantially over the years. Although the advent of operational diagnostic criteria has improved the reliability of diagnosis, this does not circumvent the fundamental problem of attempting to classify a disorder that is heterogeneous and best considered in a number of dimensions. DSM–IV and ICD–10, have virtually the same diagnostic features for a ‘clinically important’ severity of depression (termed a major depressive episode in DSM–IV or a depressive episode in ICD–10). Nevertheless their thresholds differ, with DSM–IV requiring a minimum of five out of nine symptoms (which must include depressed mood and/or anhedonia) and ICD–10 requiring four out of ten symptoms (including at least two of depressed mood, anhedonia and loss of energy). This may mean that more people may be identified as depressed using ICD–10 criteria compared with DSM–IV (Wittchen et al.

2001a), or at least that somewhat different populations are identified (Andrews et al. 2008), related to the need for only one of two key symptoms for DSM–IV but two out of three for ICD–10. These studies emphasise that, although similar, the two systems are not identical and that this is particularly apparent at the threshold taken to indicate clinical importance. The GDG considered it important to acknowledge the uncertainty inherent in our current understanding of depression and its classification, and that assuming a false categorical certainty is likely to be unhelpful and, even worse, damaging.

An important motivation has been to provide a strong steer away from only using symptom counting to make the diagnosis of depression and, by extension, to emphasise that symptom severity rating scales should not be used by themselves to make the diagnosis, although they can be an aid in assessing severity and response to treatment. To make a diagnosis of a depression requires assessment of three linked but separate factors: (a) severity, (b) duration and (c) course. Diagnosis requires a minimum of 2 weeks' duration of symptoms that includes at least one key symptom. Individual symptoms should be assessed for severity and impact on function, and be present for most of every day.

It is important to emphasise that making a diagnosis of depression does not automatically imply a specific treatment. A diagnosis is a starting point in considering the most appropriate way of helping that individual in their particular circumstances. The evidence base for treatments considered in this guideline is based primarily on randomised controlled trials (RCTs), in which standardised criteria have been used to determine entry into the trial. Patients seen clinically are rarely assessed using standardised criteria, reinforcing the need to be circumspect about an over-rigid extrapolation from RCTs to clinical practice.

Diagnosis using the three factors of severity, duration and course only provides a partial description of the individual experience of depression. People with depression vary in the pattern of symptoms they experience, their family history, personalities, premorbid difficulties (for example, sexual abuse), psychological mindedness and current relational and social problems – all of which may significantly affect outcomes. It is also common for depressed people to have a comorbid psychiatric diagnosis, such as anxiety, social phobia, panic and various personality disorders (Brown et al. 2001), and physical comorbidity. Gender and socioeconomic factors account for large variations in the population rates of depression and few studies of pharmacological, psychological or indeed other treatments for depression either control for or examine these variations. This serves to emphasise that choice of treatment is a complex process and involves negotiation and discussion with patients, and, given the current limited knowledge about which factors are associated with better antidepressant or psychotherapy response, most decisions will rely upon clinical judgement and patient preference until there is further research evidence. Trials of treatment in unclear cases may be warranted, but the uncertainty needs to be discussed with the patient and benefits from treatment carefully monitored.

The differential diagnosis of depression can be difficult; of particular concern are patients with bipolar disorder presenting with depression. The issue of differential diagnosis in this area is covered in the NICE guideline on bipolar disorder (NICE 2014).

Update 2017

2.2 Aetiology

The enormous variation in the presentation, course and outcomes of depressive illness is reflected in the breadth of theoretical explanations for its aetiology. These include processes that are genetic (Kendler and Prescott 1999), biochemical, endocrine, neurophysiological (Goodwin 2000, Malhi et al. 2005), psychological (Freud 1917, Beck 1964), and social (Brown and Harris 1978). It is important to consider these factors in understanding what predisposes to, triggers and perpetuates an episode of depression. It is also clinically apparent that features of depression itself such as loss of independence and thoughts of helplessness further compound the disability.

An emphasis upon physical and especially endocrine theories of causation has been encouraged by an observed association with some physical illnesses including diabetes, cardiac disease, hyperthyroidism, hypothyroidism, Cushing's syndrome, Addison's disease and hyperprolactinaemic amenorrhea (Cassano and Fava 2002). An association between low and very low birthweight and major depressive disorder also suggests a physical predisposition linked to intrauterine factors (Lyall et al. 2016).

Psychological theories of depression include the behavioural model in which depression results from a lack of positive reinforcement from interactions with the environment (Lewisohn et al. 1980). The cognitive model emphasises the role of cognitive distortions (biased thinking) in emotional processes (Beck 2008). The interpersonal model of depression focuses on key relationships and attachment style (Weissman et al. 2000). Some personality traits, such as neuroticism, also increase the risk of depression in the face of stressful life events (Fava and Kendler 2000). However, different personalities have different expectancies of stressful life events and some personalities have different rates of dependent life events that are directly related to their personality type, such as the end of a relationship (Hammen et al. 2000). Personality develops throughout life and certain protective characteristics may be acquired with ageing, such as self-acceptance and wisdom (Reichstadt et al. 2010).

Early life experiences such as a poor parent–child relationship, divorce, and physical and sexual abuse appear to increase a person's later vulnerability to depression (Fava and Kendler 2000). The role cannot be doubted of current social circumstances, such as poverty or unemployment, in increasing the risk of depression. Precisely how these factors interact and influence that vulnerability, however, will vary (Harris 2000). The validity of a social model of depression, in which vulnerabilities interact with stressful life events is not supported by the observation that some episodes of depression occur in the absence of a stressful event and, conversely, many such events are not followed by a depressive disorder. Lack of a confiding relationship appears to be a strong risk factor for depression (Patten 1991) and disturbances of social and leisure activities are related to severity of depression, particularly in women, and are known to persist after remission of the depressive episode (Shapira et al. 1999). Social isolation appears, in part, to account for the relationship between depression and low economic status (Bruce and Hoff 1994). While marriage appears to protect men against depression, it seems to make women more vulnerable (Weissmann 1987). Reaching old age is often associated with life events and changed social and family relationships. While older people and health care workers recognise the negative impact of loneliness, lack of social network, and reduced function, they may not recognise them as causes of depression but more an inevitable part of ageing; this can lead to negative expectations of treatment (Burroughs et al. 2006).

A family history of depressive illness accounts for around 39% of the variance of depression in both sexes (Kendler et al. 2001). Molecular genetics is making an increasing contribution to the understanding of the aetiology of depressive disorders, adding to the work in genetic epidemiology. Evidence for the interaction of genes and environment in conferring vulnerability to depression is suggested by the finding of a polymorphism in the serotonin transporter gene of people with a greater tendency to depression in the face of negative life events (Caspi et al. 2003), although this association remains controversial. It has been suggested that genetic factors may be less important when the onset of depression is late in life (Baldwin 2012). Genetic and psychological theories are now being linked. For instance, a hypersensitive amygdala is known to be associated with both a genetic polymorphism and a pattern of negative cognitive biases and dysfunctional beliefs, all of which constitute risk factors for depression (Beck 2008). Further, the combination of a hyperactive amygdala and hypoactive prefrontal cortex is associated with diminished cognitive appraisal and the occurrence of depression.

Advances in neuroimaging have reinforced the idea of depression as a disorder of brain structure and function (Drevets et al. 2008) and in older people, the presence of cerebral

Update 2017

white matter changes on magnetic resonance imaging predicts the onset of depression (Teodorczuk et al. 2010). The causes of late-life depression are thought to differ from depression in younger adults, especially in cases with onset after 50 years of age, which have greater neuropsychological abnormalities such as executive dysfunction (Gansler et al. 2015). There has been much interest in recent years in a possible association between cardiovascular risk factors and depression ('vascular depression') in later life but with inconsistent findings on the strength of any association and the direction of causality. A systematic review of relevant studies suggests that depression is associated with active cardiovascular disease, diabetes and stroke, but not with hypertension, smoking, and dyslipidaemia (Valkonova and Ebmeier 2013). There is a complex aetiological and clinical interplay between late-life depression, cognitive impairment, and dementia (Baldwin 2012).

Health care workers should be aware of the negative impact on mood of discrimination experienced by people from black and minority ethnic communities and work to ensure equal access to people from all ethnic backgrounds (Department of Health 2005). In England and Wales, there is diversity of minority ethnic communities including Irish, African-Caribbean and Asian. Social disadvantage and real or perceived prejudice may contribute to the onset of depression, and delays in help-seeking or miscommunication with professionals may perpetuate problems (Craig and Bhugra 2012).

People from the lesbian, gay, bisexual and transgender (LGBT) communities may be vulnerable to depression at certain times (http://pinkthearpy.mobi). There are few epidemiological studies of depressive disorders in the LGBT communities. However, while there appears to be no difference in the prevalence of depressive symptoms between homosexual and heterosexual people of stable sexual orientation, changes in sexual identity and disclosure of sexual orientation or gender identity are associated with a higher incidence of depression (Everett, 2015, Nuttbrock et al. 2011; Pachankis et al. 2015). Older lesbian, gay and bisexual people may also face mental health problems associated with isolation and a reluctance to disclose their orientation to health professionals (Guasp et al. 2010).

Depressive illness is frequently a long-term condition of fluctuating intensity. The range of factors known to be associated with chronic depression is large. Among the most important of these are a family history of depression, comorbid anxiety disorder, substance abuse, dependent and avoidant personality disorders, advancing age and low income (Blanco et al. 2010). Clinicians need to be aware of the substantial unmet treatment needs in chronic depression and consider the scope for intervention with these known prognostic factors.

Update 2017

2.3 Daily life: family and relationships

Depression is related to family and couple stress and conflict in a bi-directional way: depression is both caused by and is itself the cause of difficult family relationships (Davila, Karney, Hall and Bradbury, 2003), however there is evidence that distressed couple and marital relationships have a greater impact on the likelihood of major depression than distress in relationships with other family members and close friends (Whisman, Sheldon and Goering 2000). Whisman calculated that individuals in couple relationships that were distressed were 3 times more likely to have a mood disorder than individuals in a relationship that was not distressed, and Whisman and Uebelacker (2003) estimate that up to 30% of severe depressive episodes could be prevented if the couple relationship was improved. Depression was linked to the length of the couple relationship and the severity of conflict by Kouros and colleagues (Kouros, Papp, & Cummings 2008).

In addition, there are clear links between family disagreements (usually defined in terms of the quality of the couple relationship), somatic symptoms and depression, with a 23-year study by Bi et al (Bi, Breland, Moos & Cronkite, 2015) confirming that in families where there is depression there is a greater amount of disagreement and somatic symptoms than in non-depressed families. Life satisfaction and relationship adjustment mutually influence each

other, with a greater influence of relationship adjustment on life satisfaction for women according to Be and colleagues (Be, Whisman and Uebelacker, 2013).

Segrin (2000) has reviewed the relationship between poor social skills and depression and concluded that the evidence is equivocal in relation to directionality, but that it confirms that depression and poor social skills are concomitant. Choi and Marks (2008), on the other hand, concluded that marital difficulties led directly to both depression and functional impairment. This suggests that, if difficulties in relating are not addressed, depression may not lift as much as it might have done.

The London Depression study (Leff et al. 2000) indicated that depression and critical comments from partners are linked, and that couples in this study preferred therapy to antidepressants, with only 15% of participants in the couple therapy arm dropping out of treatment as compared to 56.8% of those in the medication arm.

The Teo et al. 10-year follow-up study of people with social strain and poor quality of relationships (Teo, Choi and Valenstein 2013) showed that social isolation alone was not predictive of future incidents of depression, whereas poor quality of relationships with spouses, and to a lesser extent with family members – but not with friends – was predictive of future incidents of depression 10 years later. People with a lot of relationship strain were more than twice as likely to have an episode of major depression as those with little relationship strain. This effect occurred even if there had not been a prior history of depression, though for this group difficulty in relation to a spouse or partner and not family members or friends was significantly associated with future depression. This finding echoes other studies such as Beach et al. (Beach, Katz, Kim & Brody 2003) and the work of Cano and O'Leary (Cano & O'Leary 2000) showing that humiliating events for women in marital relationships (infidelities and threats of separation) are 6 times more likely to result in an episode of major depressive disorder than in a control group where there was not such humiliation. Beach and colleagues (Beach et al. 2004) have shown that incidents of physical aggression aimed at wives in heterosexual relationships also increase the risk of subsequent depression.

Update 2017

Foran et al (Foran, Whisman and Beach, 2015) pointed out how the outcomes of individual psychotherapy and psychopharmacological treatment for depression are detrimentally affected by relationship distress (Denton et al. 2010) and that relationship distress also predicts relapse including for people who have been successfully treated for depression (whether by individual psychotherapy or psychopharmacological treatments) (Whisman 2001).

There is also evidence that treating relationship distress reduces subsequent health service usage by 22% (Law and Crane 2000), with higher users (defined as having four or more visits within 6 months) reducing their usage of urgent care by 78% after receiving conjoint therapy (Law, Crane and Berge 2003), underlining the importance of attending to the close relationships that people experiencing episodes of depression have.

2.4 Treatment and management of depression in the National Health Service

2.4.1 Detection, recognition and referral in primary care

Of the 130 cases of depression (including mild cases) per 1000 people per year, only 80 will consult their GP. The most common reasons given for reluctance to contact the family doctor include: not thinking anyone could help (28%); feeling it was a problem one should be able to cope with (28%); not thinking it was necessary to contact a doctor (17%); thinking the problem would get better by itself (15%); feeling too embarrassed to discuss it with anyone (13%); and being afraid of the consequences (for example, treatment, tests, hospitalisation,

being sectioned; 10%) (Meltzer et al. 2000). The stigma associated with depression cannot be ignored in this context (Priest et al. 1996).

Initial recognition

Of the 80 depressed people per 1000 who do consult their GP, 49 are not recognised as depressed on the first visit, mainly because most of them are consulting for a somatic symptom and do not consider themselves mentally unwell, despite the presence of symptoms of depression (Kisely et al. 1995). Depression is much more likely to be recognised when people present with psychosocial symptoms rather than somatic symptoms (Kirmayer et al. 1993, Tylee et al. 1995). Those people who go unrecognised on a single occasion also have milder illnesses (Goldberg et al. 1998), and GPs are better at recognising moderate to severe depression for which the evidence of benefit from treatment is stronger (Thompson et al. 2001). In addition, longitudinal research suggests most patients who are unrecognised on a single occasion are subsequently recognised and treated, although a minority of cases can remain undetected and unwell for years (Kessler et al. 2002).

GPs are immensely variable in their ability to recognise depressive illnesses, with some recognising virtually all the patients found to be depressed at independent research interview, and others recognising very few (Goldberg & Huxley 1992, Üstün and Sartorius 1995). The communication skills of the GP 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 facilitating detection (Goldberg and Bridges 1988, Goldberg et al. 1993).

Attempts to improve the rate of recognition of depression by GPs using guidelines, lectures and discussion groups have not improved recognition or outcomes (Thompson et al. 2000, Kendrick et al., 2001), although similar interventions combined with skills training may improve detection and outcomes in terms of symptoms and level of functioning (Tiemens et al. 1999, Ostler et al. 2001). However, the inference that these health gains are the result of improved detection and better access to specific treatments, while having face validity, has been contested. For example, Ormel and colleagues (1990) suggested that the benefits of recognition of common mental disorders could not be attributed entirely to specific mental health treatments. Other factors, such as acknowledgement of distress, reinterpretation of symptoms, and providing hope and social support, were suggested to contribute to better patient outcomes.

This view gained confirmation from a Dutch study in which providing skills training for GPs did not improve detection, but did improve outcomes. Moreover, about half of the observed improvement in patient outcomes was mediated by the combined improvements in process of care. In combination with the strong mediating effect of empathy and psychoeducation they suggested that other, probably also non-specific, aspects of the process of care might be responsible for the training effect on symptoms and disability (Van Os et al. 2004). In addition, the communication skills needed by GPs can be learned and incorporated into routine practice (Gask et al. 1988, Roter et al. 1995), although interventions sometimes fail to impact on patient outcomes, despite changes in clinician behaviour (Gask et al. 2004).

Screening and case finding

The fact that common mental health disorders often go undiagnosed among primary care attenders has led to suggestions that clinicians should systematically screen for hidden disorders. However, general screening has not been shown to improve patient outcomes (Gilbody et al. 2008), and is currently not recommended in most countries, including the UK (Gilbody et al. 2006). Instead, targeted case finding, which involves screening a smaller group of people known to be at higher risk based on the presence of particular risk factors, may be a more useful method of improving the recognition of depression in primary care (see Chapter 6). Furthermore, research suggests improved detection alone does not improve

Update 2017

patient outcomes in the absence of improved treatments being provided for those detected (Gilbody et al. 2003).

Referral

Of those people that are recognised as depressed, most are treated in primary care and only about one in four or five are referred to psychological therapies or secondary mental health services (Kendrick et al. 2009). There is considerable variation among individual GPs in their referral rates to mental health services, but those seen by specialist services are a highly selected group – they are skewed towards those who do not respond to antidepressants, people with more severe illnesses, single women and those below 35 years of age (Goldberg and Huxley 1980).

In summary, those with more severe disorders, and those presenting with psychological symptoms, are especially likely to be recognised as depressed while those presenting with somatic symptoms for which no obvious cause can be found are less likely to be recognised. The evidence suggests that these very undesirable circumstances, in which large numbers of people each year experience depression, with all of the attendant negative personal and social consequences, could be changed. With 50% of people with depression never consulting a doctor, 95% never entering secondary mental health services, and many more whose depression goes unrecognised and untreated, this is clearly a problem for primary care.

2.4.2 Assessment and co-ordination of care

Given the low detection and recognition rates, it is essential that primary care and mental health practitioners have the required skills to assess people with depression, their social circumstances and relationships, and the risk they may pose to themselves and others. This is especially important in view of the fact that depression is associated with an increased suicide rate, a strong tendency for recurrence, and high personal and social costs. The effective assessment of a patient, including risk assessment and the subsequent co-ordination of their care (through the use of the Care Programme Approach [CPA] in secondary care services), is highly likely to improve outcomes and should, therefore, be comprehensive.

Update 2017

2.4.3 Aim, and non-specific effects, of treatment and the placebo

The aim of intervention is to restore health through the relief of symptoms and restoration of function and, in the longer term, to prevent relapse. Where possible, the key goal of an intervention should be complete relief of symptoms (remission), which is associated with better functioning and a lower likelihood of relapse (Kennedy and Foy, 2005). It may not always be possible to achieve remission, but it is usually possible to improve symptoms and functioning to an important degree. For this reason the GC examined a range of outcomes (where available), including response, remission, change in symptoms and relapse. The relative importance of these depends on many factors, including the severity of depression, the degree of impairment to everyday functioning experienced and the patient's psychiatric history. Among those seeking treatment for depression, those put on waiting lists do improve steadily with time. Posternak and Miller (2001) studied 221 patients assigned to waiting lists in 19 treatment trials of specific interventions and found that 20% improved within 4 to 8 weeks, and 50% improved within 6 months. They estimated that 60% of responders to placebo and 30% of responders to antidepressants may experience spontaneous resolution of symptoms (if untreated). An earlier study by Coryell and colleagues (1994) followed up 114 patients with untreated depression for 6 months: the mean duration of an episode was 6 months, with 50% remission in 25 weeks. It should be noted that there is a high relapse rate associated with depression (see Section 2.1.2, above).

Despite their greater severity and other differences, Furukawa and colleagues (2000) showed that patients treated by psychiatrists with antidepressants showed greater improvements than untreated patients: the median time to recovery was 3 months, with 26% recovering in 1 month, 63% in 6 months; 85% in 1 year, and 88% in 2 years.

Although there is insufficient space here to allow proper discussion, it should be noted that non-specific/placebo effects apply not only to treatment with medication but also to other treatments. Studies comparing any treatment with a waiting list control or treatment as usual (TAU) in which there is minimal intervention are therefore difficult to interpret and improvements could simply be due to the increased support, engagement and monitoring that the intervention involves.

The placebo effect in trials of psychiatric drugs is often so large that specific pharmacological effects can be hard to identify, especially when given to people who fall into one of the larger, more heterogeneous diagnostic categories. There can also be suspicion of publication bias, especially with regard to drug company funded trials (Lexchin et al. 2003, Melander et al. 2003). A meta-analysis by Kirsch et al. (2008) of all data submitted to the US Food and Drug Administration (FDA) for the licensing of new antidepressants was controversial in suggesting that the overall effect of new-generation drugs including the SSRIs and venlafaxine was below the NICE recommended criteria for clinical relevance. They suggested that efficacy reached clinical relevance only in trials involving the most extremely depressed patients, and that this was due to a decrease in the response to placebo rather than an increase in the response to medication. A subsequent meta-analysis of similar data by Fournier et al. (2010) also suggested that, while for patients with very severe depression the benefit of medications over placebo is substantial it may be minimal or non-existent in patients with mild or moderate symptoms. Turner et al. (2008) found that selective publication of drug company funded trials with positive findings led to an overestimation of the benefits of active drugs over placebo. A re-analysis of the FDA data by Fountoulakis and Möller (2011) suggested however that Kirsch et al.'s (2008) meta-analysis suffered from selective reporting of the results and that their conclusions were unjustified and overemphasised. The authors suggested that, although a large percentage of the placebo response is due to expectancy, this is not true for the response to the active drug and the effects are not additive. In other words the contribution of the biochemical effect of the drug is always present and is unrelated to depression severity, while the contribution of the psychological placebo effect varies – it contributes a greater proportion of the effect in mild depression than in severe depression (Fountoulakis and Möller 2011).

Antidepressants (or other) treatments for depression may therefore offer little or no advantage, on average, over placebo for patients with subthreshold depressive symptoms or mild depression, who often improve spontaneously or who respond well to non-specific measures such as support and monitoring. The evidence does however support the efficacy of specific treatments with more severe depression and in those with depression that persists over time.

At present it is not possible to clearly identify people with depression who will respond to the specific aspects of a treatment as opposed to the non-specific effects associated with having a treatment. Weimer et al. (2015) reviewed 31 meta-analyses and systematic reviews of more than 500 randomised placebo-controlled trials across a range of psychiatric conditions including depression, to identify factors associated with an increased placebo response. Of 20 factors discussed, only three were often linked to high placebo responses: low baseline severity of symptoms, more recent trials, and unbalanced randomisation (more patients randomly assigned to drug than placebo). Laboratory studies with psychological, brain, and genetic approaches had not successfully identified predictors of placebo responses and the authors concluded that predictors of the placebo response are still to be discovered.

Update 2017

2.4.41 Pharmacological treatments

The mainstay of the pharmacological treatment of depression for the last 50 or more years has been antidepressants. Tricyclic antidepressants (TCAs) were introduced in the 1950s, the first being imipramine (Kuhn 1958). The mode of action of this class of drug, thought to be responsible for their mood-elevating properties, is their ability to block the synaptic reuptake of monoamines, including noradrenaline (NA), 5-hydroxytryptymine (5HT) and dopamine (DA). In fact, the TCAs predominantly affect the reuptake of NA and 5HT rather than DA (Mindham 1982). The antidepressant properties of monoamine-oxidase inhibitors (MAOIs) were discovered by chance in the 1950s, in parallel with TCAs.

Although the introduction of the TCAs was welcome, given the prior lack of specific treatments for people with depression, the adverse effects resulting from their ability to influence anticholinergic, histaminergic and other receptor systems compromised their acceptability. Moreover, overdose with TCAs (with the sole exception of lofepramine) carries a high mortality and morbidity. This is obviously particularly problematic in the treatment of people with suicidal intentions.

Because of the side-effect profile of TCAs and related drugs and their toxicity in overdose, new classes of antidepressants were developed, including: selective serotonin reuptake inhibitors (SSRIs), such as fluoxetine and sertraline; drugs chemically related to but pharmacologically different from the TCAs, such as trazodone; and a range of other chemically unrelated antidepressants, including mirtazapine and agomelatine. Their effects and adverse effects vary considerably, although their mood-elevating effects are again thought to be mediated through increasing intra-synaptic levels of monoamines, some primarily affecting NA, some 5HT and others affecting both to varying degrees and in different ways. The most recently introduced drugs may have somewhat different modes of action. Agomelatine, uniquely, is a melatonin agonist and vortioxetine is an SSRI with additional activity at 5HT1A and 5HT7 receptors. Despite somewhat different pharmacological effects, all antidepressants seem to share ‘downstream’ effects on inflammatory markers and brain-derived neurotropic factor. There is also evidence to support a cognitive neuropsychological model of therapeutic action whereby antidepressants are thought to remediate negative biases in emotional processing from an early stage of treatment (Walsh and Harmer 2015).

Update 2017

Other drugs used either alone or in combination with antidepressants include lithium and some atypical antipsychotics, although the use of these drugs is usually reserved for people with refractory or psychotic depressions.

There is very preliminary evidence that pharmacogenetic variations may affect the efficacy and tolerability of antidepressant drugs. It is likely that future research on this topic will lead to the development of clinically meaningful pharmacogenetic markers, but at the moment the data are insufficient to make recommendations.

2.4.59 Psychological treatments

Numerous theories and methods for the psychological treatment of depression have been elaborated and championed over the last 40 years since the pioneering efficacy research on cognitive and behavioural approaches (Beck et al. 1979). There is a growing emphasis upon the evidence base and the specific adaptation of psychological treatments for people with depression, although systematic research into what works for whom is still evolving (Roth and Fonagy, 2005, Cartwright and Munro 2010). Nonetheless, a range of psychological and psychosocial interventions for depression have been shown to relieve the symptoms of the condition, with growing evidence that psychological therapies can help people recover from depression in the longer-term (NICE 2009).

Psychological treatments for depression currently claiming efficacy in the treatment of people with depressive illnesses and reviewed for this guideline include: cognitive behavioural

therapy (CBT); behavioural activation; interpersonal therapy (IPT); problem-solving therapy; counselling; psychodynamic psychotherapy; and couples therapy. Psychological treatments generally have more widespread acceptance than medication from service users (Priest et al. 1996, van Schaik et al. 2004) with a recent meta-analysis suggesting a 3-fold preference for psychological treatment (McHugh et al. 2013). It is also increasingly recognised that individuals wish to have a choice of psychological treatment options, and that the provision of such choice may improve treatment engagement and outcome (Kocsis et al. 2009; Swift and Callahan 2009).

This guideline distinguishes between high-intensity and low-intensity psychological interventions. High-intensity interventions are typically psychological therapies such as CBT, IPT, psychodynamic, or couples therapy provided by a therapist face-to-face over an extended duration of sessions. Within these therapies, formulation of each individual presentation informs treatment options and therapists have flexibility in treatment delivery. In contrast, low-intensity interventions typically involve guided written or audio-recorded self-help materials or computerised or internet-delivered CBT, where a practitioner facilitates and supports the use of these materials, or group work. Low-intensity interventions are typically briefer and more generic, enabling a greater volume of people with depression to be seen per practitioner. Training to deliver high-intensity therapies typically involves an extensive period of supervised practice in a specific evidence-based model for already qualified mental health professionals, whereas training for low-intensity interventions uses a briefer structured protocol-led approach including specific assessments of competency, as in the training of Psychological Wellbeing Practitioners (PWPs). Because both high- and low-intensity therapies have demonstrated efficacy, a Stepped/Matched Care model was recommended by the previous guideline (NICE 2009), in which interventions demanding less resources are offered first, where clinically appropriate (Bower and Gilbody 2005).

Since the publication of the previous guideline (2009), the provision of psychological therapies has been significantly expanded by the Improving Access to Psychological Treatments (IAPT) programme. This has involved the national roll-out in England of primary care delivery sites to provide evidence-based NICE-recommended high- and low-intensity psychological interventions, predominantly CBT and counselling. The use of high- and low-intensity interventions within a stepped care framework has enabled many more people to access and complete psychological treatment on the NHS than previously (over 500,000 per year according to national figures [HSCIC 2015]) with over 2/3rds seen within 4 weeks. Nonetheless, there is considerable scope for improvement, as IAPT still only meets an estimated 15% of the need for common mental health problems in adults, many people cannot access their preferred psychological treatment, attrition is high, and there is not equity of access, especially for black and minority ethnic (BME) groups and older people (HSCIC 2015). In addition, there remain commissioning issues concerning capacity, particularly for individuals to receive an adequate number of high-intensity intervention sessions, and workforce training, where therapists cannot access the required training to deliver specific evidence-based psychological therapies (NAPT 2013).

Update 2017

2.4.6 Physical treatments

Aside from pharmacological treatments, there is a diverse range of physical treatments sometimes used in the management of depressive illness. Of these, electroconvulsive therapy (ECT; electroplexy) is the most established; other treatments lack strong evidence of efficacy. These treatments may be chosen due to patient preference or for the treatment of specific sub-types of depression. They may also be used when pharmacotherapy or psychological treatment are unsuccessful, or in conjunction with them.

Electroconvulsive therapy

Electroconvulsive therapy is widely available in England and Wales where its use is regulated by the Royal College of Psychiatrists ECT Accreditation Service (Hodge and Buley

2014). It originated as a treatment for mental illness in the 1930s after the observation that chemically-induced seizures improved the outcome of catatonic schizophrenia; later, electrical induction of seizures was developed (Shorter et al. 2007). The mechanisms by which ECT restores normal brain function remain unclear, but they are thought to include effects on cerebral blood flow, cerebral metabolism, nerve growth and plasticity, neurotransmitter pathways, and neuroendocrine systems (Anderson and Fergusson 2013). Over recent years, the mode of administration of ECT has been significantly refined to maximise efficacy and limit side-effects. Electroencephalogram (EEG) monitoring of treatment is now standard practice and, to achieve efficacy, seizures are induced at 1.5 to 2.5 times seizure threshold with bitemporal electrode placement or at 6 to 8 times seizure threshold with unilateral placement (Fink 2012). Most treatment courses are between 6 and 20 sessions. The efficacy of ECT probably exceeds that of pharmacotherapy (UK ECT Review Group 2003). Occasionally, continuation and maintenance ECT is recommended when the risk of relapse or recurrence is very high. Short-term cognitive impairment is commonly reported after ECT and longer term impairment of autobiographical memory may also be a consequence (Freeman 2013) which is a particular concern with older patients. Unilateral electrode placement is thought less likely to induce cognitive side-effects but is less efficacious. The use of shorter electrical pulse widths over recent years has helped to limit cognitive side-effects and there is now interest in the use of even shorter (ultra-brief) pulses (Tor et al. 2015). Due to its invasive nature, the risk of cognitive impairment and the need for general anaesthesia, ECT is usually used for the treatment of severe, high risk depression or following unsuccessful treatment with pharmacotherapy. Despite it now being administered in modern, regulated facilities ECT still attracts a negative public image.

Other brain stimulation therapies

Other electrical techniques for the treatment of depression that modulate brain activity without inducing seizures include repetitive transcranial magnetic stimulation (rTMS), transcranial direct current stimulation (tDCS), vagus nerve stimulation (VNS), and deep brain stimulation (DBS) (Brunoni et al. 2010). rTMS, tDCS and VNS for the treatment of depression are outside the scope of this guideline and are addressed in other NICE guidance (NICE 2015-1, NICE 2015-2; NICE 2009). Of the three techniques, the greatest evidence base exists for rTMS. VNS and DBS are both invasive techniques. A recent randomised controlled trial of DBS applied to the ventral capsule and ventral striatum failed to show superiority of active over sham stimulation (Dougherty et al. 2015).

Update 2017

Phototherapy

Descriptions of the benefits on mood of light exposure go back at least to the second century and artificial bright light treatment (phototherapy) has been studied in the treatment of depression since the description of seasonal affective disorder in the 1980s (Cowen, 2012). It is thought to act by advancing endogenous circadian rhythms (Lewy et al. 1987). Therapy is usually delivered using a light box made up of fluorescent tubes. Variable treatment parameters include light intensity (measured in lux) and frequency and duration of exposure. Artificial light therapy is usually well tolerated but side-effects include headache and eye irritation.

Acupuncture

Traditional acupuncture uses needle puncture of the skin over specific designated anatomical points in the treatment of pain and other conditions, including depression. Laser acupuncture is a newer technique that avoids puncturing the skin (Quah-Smith et al. 2013). Some studies suggest that acupuncture may augment the effect of antidepressant treatment (Chan et al. 2015).

Aromatherapy

Aromatherapy has been used in the treatment of a range of medical conditions, including depression. It involves the application of plant-derived oils via massage into the skin or inhalation from infusers. Due to a small number of studies of poor quality, its efficacy in depression is unclear (Lee et al. 2012).

2.4.76 Service-level and other interventions

Given the complexity of healthcare organisations, and the variation in the way care is delivered (inpatient, outpatient, day hospital, community teams, and so on), choosing the right service configuration for the delivery of care to specific groups of people has gained increasing interest with regard to both policy (for example, see Department of Health, 1999), and research (for example, evaluating day hospital treatment, Marshall et al., 2001). Research using RCT designs has a number of difficulties; for example, using comparators such as 'standard care' in the US make the results difficult to generalise or apply to countries with very different types of 'standard care'.

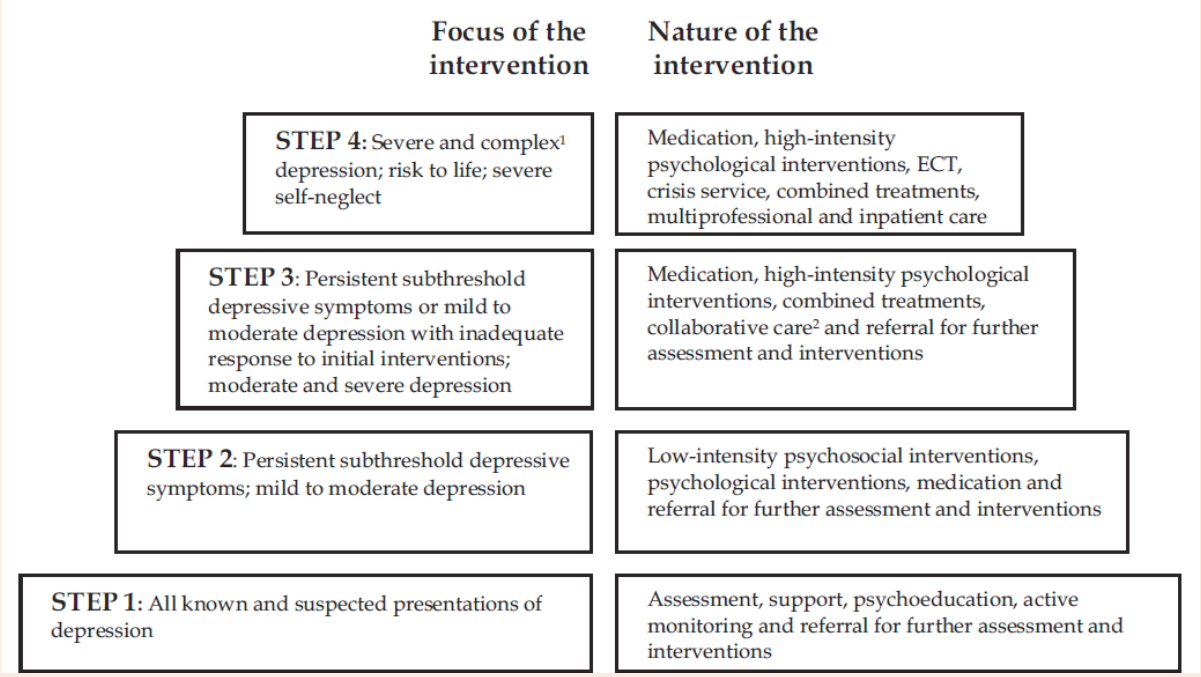
Service-level interventions considered for review in this guideline include: organisational developments, crisis teams, day hospital care, non-statutory support and other social supports. Other types of interventions reviewed for this guideline include: physical activity programmes, guided self-help, computerised cognitive behavioural therapy (CCBT) and screening.

2.4.80 Delivery of care

In Figure 1, a 'stepped-care' model is developed that draws attention to the different needs that depressed individuals have – depending on the characteristics of their depression and their personal and social circumstances – and the responses that are required from services. Stepped care provides a framework in which to organise the provision of services supporting patients, carers and healthcare professionals in identifying and accessing the most effective interventions.

Update 2017

Figure 1: The stepped-care model



Notes: 1 'Complex depression' here includes depression that shows an inadequate response to multiple treatments, is complicated by psychotic symptoms, and/or is associated with significant psychiatric comorbidity or psychosocial factors.2 Only for depression where the person also has a chronic physical health problem and associated functional impairment (see NICE 2009c).

Of those people whom primary healthcare professionals recognise as having depression, some prefer to avoid medical interventions and others will improve in any case without them. Thus, in depression of only mild severity, many GPs prefer an 'active monitoring' approach, which can be accompanied by general advice on such matters as restoring natural sleep rhythms and getting more structure into the day. However, other people prefer to accept, or indeed require, medical, psychological or social interventions, and these patients are therefore offered more complex interventions. Various interventions are effective, delivered by a range of workers in primary care.

Treatment of depression in primary care, however, often falls short of optimal guideline recommended practice (Donoghue & Tylee, 1996) and outcomes are correspondingly below what is possible (Rost et al., 1995). As we have seen, only about one in five of the patients at this level will need referral to a mental healthcare professional, the main indications being failure of the depression to respond to treatment offered in primary care, incomplete response or frequent recurrences of depression. Those patients who are actively suicidal or whose depression has psychotic features will need specialist referral.

Finally, there are a few patients who will need admission to an inpatient psychiatric bed. Here, they can receive 24-hour care and various special interventions.

2.58 The economic cost

Update 2017

Depression places a significant burden on individuals and their carers, health services and communities worldwide. According to the Global Burden of Diseases, Injuries, and Risk Factors Study 2010, major depression was the leading cause of disability among mental health and behavioural disorders worldwide, and the 11th single leading cause of disability among 291 diseases and injuries, accounting for 2.5% of disability-adjusted life years (DALYs) in 2010 (Murray et al. 2012); in terms of number of years lived with disability (YLD), major depression ranked 2nd single leading cause, accounting for 9.6% of YLD globally (Vos et al. 2012). In Western Europe, major depression was found to be the 4th single leading cause of DALYs and the 2nd single leading cause of YLD among all causes. The global burden of disease caused by unipolar depression (including major depression and dysthymia) increased by 38% from 1990 to 2010. (Murray et al. 2012).

A UK study estimated the total cost of depression in adults in England in 2000 (Thomas & Morris 2003). A prevalence-based approach was used by applying rates of depression from Office of National Statistics (ONS) data to population data for England in 2000. The study measured the direct treatment costs of depression, including primary and secondary care costs as well as indirect costs of lost working days (morbidity) and lost life-years (mortality). The direct treatment costs were estimated at £370 million, of which 84% were attributable to antidepressant medication, 7% to inpatient care, 6% to outpatient and day care and 3% to primary care services. However, the indirect costs of depression were estimated to be far greater: total morbidity costs were more than £8 billion and mortality costs reached £562 million. In comparison with the findings of earlier UK-based cost-of-illness studies, direct treatment costs shifted from hospital admissions (including specialised mental institutions) towards medication, reflecting changes in patterns of care over time away from expensive inpatient care to relatively less expensive outpatient-based care but also greater usage of more expensive, patented antidepressants.

More recently, McCrone and colleagues (2008) estimated the total mental health expenditure in England for 20 years (2007-2026). The study combined prevalence of the most major mental disorders, taken from the Psychiatric Morbidity Survey 2000 (Singleton et al 2001),

with population estimates from 2007 through to 2026. It was estimated that in 2007 there were 1.24 million people with depression in England, and this number was projected to rise by 17% to 1.45 million by 2026 due to demographic changes. Based on these figures, the authors estimated the total service costs for depression in England for 2007 at £1.7 billion. This cost accounted for prescribed drugs (1%), GP care (9%), inpatient care (10% psychiatric and 17% non-psychiatric), other NHS non-inpatient services (33%), residential care (10%), other social service costs (15%) and other costs (5%). Including the cost of lost employment in terms of workplace absenteeism resulted in the total cost of depression reaching £7.5 billion. By 2026 these figures were projected to be £3 billion for total service costs and £12.2 billion if lost employment was also considered. In contrast to the study by Thomas and Morris (2003), antidepressant medication accounted for only 1% of total service costs whilst secondary care accounted for over 50% of these costs. However, in both studies, lost employment was by far the driver of the total cost, contributing to the estimated figure by more than 75%.

Sobocki and colleagues (2006) estimated that in 28 European countries with a total population of 466 million, at least 21 million were affected by depression. The authors reported an estimated total annual cost of depression in Europe of €118 billion in 2004, corresponding to a cost of €253 per inhabitant. Direct healthcare costs reached €42 billion, comprising €22 billion outpatient care costs, €9 billion drug costs, and €10 billion hospitalisation costs. Indirect costs due to morbidity and mortality were estimated at €76 billion. Based on these figures, the authors concluded that depression is the most costly brain disorder in Europe, accounting for 33% of the total cost of brain disorders.

Sanderson and colleagues (2003) estimated the total direct mental healthcare cost of depression in Australia at $484 million in 2003 or $1,239 per treated case (1997–98, Australian dollars); the respective cost for dysthymia reached $71 million or $1779 per treated case. The authors estimated that if evidence-based, optimal treatment was implemented, the total direct mental healthcare cost of depression and dysthymia would fall at $341 million ($874 per treated case) and $29 million ($721 per treated case), respectively. In the US, Greenberg and colleagues (2015) estimated the total cost of major depression using national survey and administrative claims data. This cost was reported to reach $210.5 billion in 2010, comprising 45% direct healthcare costs, 5% suicide-related costs, and 50% indirect productivity losses. In Japan, the total cost of depression in 2008 was estimated to reach $11 billion, with $1.6 billion accounting for direct medical costs, $2.5 billion attributable to depression-related suicide costs, and $6.9 billion relating to lost productivity (Okumura and Higuchi 2011).

The costs of minor depression are not negligible. Cuijpers and colleagues (2007) conducted a large population-based study to estimate the costs of minor depression in the Netherlands. Excess costs, i.e. the costs of the disorder over and above the costs attributable to other illnesses, were estimated with the help of regression analysis. The authors found that the annual excess cost of minor depression was $2141 per person (2003 US dollars), while the respective cost of major depression was $3313. This cost included direct medical and non-medical costs as well as productivity losses. Using these estimates and the baseline cost attributable to other illnesses of $1023 per person, the authors estimated the total annual cost of minor depression at $160 million per 1 million inhabitants in the Netherlands, which was comparable to the respective total annual cost of £192 million estimated for major depression.

Non-adherence to antidepressant treatment leads, as expected, to increased symptom severity, decreased response and remission rates, increased risk of relapse, and higher rates of healthcare utilisation, leading to increased healthcare costs (Ho et al. 2016). Failure of treatment (due to either non-adherence or to inefficacy of treatment) considerably increases the cost of depression. Evidence from the UK (Byford et al. 2011), Sweden (Sobocki et al. 2006, von Knorring et al. 2006) and the US (Dennehy et al. 2015) suggests that non-remitters or non-responders to treatment have more contact with primary care and

Update 2017

secondary outpatient care services and a higher number of sick leave days compared to remitters, translating into a significantly higher cost compared with people with depression achieving remission following treatment.

Treatment-resistant depression appears to contribute significantly to the total cost of depression: a review of 62 studies on 59,462 people with depression reported an increase in the annual healthcare and lost productivity cost of $5,481 and $4,048, respectively, per person with treatment-resistant depression in comparison to a person with treatment-responsive depression in 2012 US dollar prices (Mrazek et al. 2014). Using these figures and prevalence of treatment-resistant depression of 12-20% among all adults with depression in the US (estimated to reach 16 million people), the authors reported an annual societal cost of $18-$30 billion attributable to treatment-resistant depression in the US, pushing up the total societal cost of major depression in the US to a total of $188-$200 billion, which is broadly consistent with the figure quoted by Greenberg and colleagues (2015).

One of the key findings from the cost-of-illness literature is that the indirect costs of depression are by far the most significant driver of the total costs of depression, being substantially higher than the health service costs. Other intangible costs of depression include the impact on the quality of life of adults with depression as well as their carers and families.

The findings of the cost-of-illness studies globally suggest that depression imposes a significant burden on individuals and their carers, family members, the healthcare system and also the broader economy through lost productivity and workplace absenteeism. Furthermore, it is anticipated that these costs will continue to rise significantly in future years. Therefore, it is important that available healthcare resources are used efficiently to maximise the benefits for people with depression, their carers and family, and the wider society.

Update 2017

3 Methods used to develop this guideline

3.1 Overview

The development of this guideline followed *Developing NICE guidelines: the manual*. A team of health care professionals, lay representatives and technical experts known as the Guideline Committee (GC), with support from the NCCMH and NGA staff, undertook the development of a person-centred, evidence-based guideline. There are 7 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 GC 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 GC adopted informal methods 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 GC.

Update 2017

3.2 The scope

Topics are referred by NHS England and the letter of referral defines the remit, which defines the main areas to be covered. The NCCMH developed a scope for the guideline based on the remit (see Appendix A). 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 GC
- encourage applications for GC membership.

The draft scope was subject to consultation with registered stakeholders over a 4-week period. During the consultation period, the scope was posted on the NICE website. 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 Committee

During the consultation phase, members of the GC were appointed by an open recruitment process. GC membership consisted of: professionals in psychiatry, clinical psychology, nursing and general practice; academic experts in psychiatry and psychology; commissioning managers; and carers and representatives from service user and carer organisations. The guideline development process was supported by staff from the NCCMH and the NGA, who undertook the clinical and health economic literature searches, reviewed and presented the evidence to the GC, managed the process, and contributed to drafting the guideline.

3.3.1 Guideline Committee meetings

There were 14 GC meetings held between June 2015 and March 2017. During each day-long GC meeting, in a plenary session, review questions and clinical and economic evidence were reviewed and assessed, and recommendations formulated. At each meeting, all GC members declared any potential conflicts of interest (see Appendix B), and service user and carer concerns were routinely discussed as a standing agenda item.

3.3.2 Service users and carers

Individuals with direct experience of services gave an integral service-user focus to the GC and the guideline. They contributed as full GC members to writing the review questions, providing advice on outcomes most relevant to service users and carers, 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 GC. They contributed to writing the guideline's introduction and identified recommendations from the service user and carer perspective.

3.3.3 Expert advisers

Expert advisers, who had specific expertise in one or more aspects of treatment and management relevant to the guideline, assisted the GC, commenting on specific aspects of the developing guideline. Appendix C lists those who agreed to act as expert advisers.

3.3.4 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 GC 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 GC 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 GC could be provided with full access to the complete trial report. Appendix E lists researchers who were contacted.

3.4 Review protocols

Review questions drafted during the scoping phase were discussed by the GC at the first few meetings and amended as necessary. The review questions were used as the starting point

Update 2017

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 F.

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?

For each topic, addressed by one or more review questions, a review protocol was drafted by the technical team using a standardised template (based on PROSPERO), reviewed and agreed by the GC (all protocols are included in Appendix F).

To help facilitate the literature review, a note was made of the best study design type to answer each question. There are 4 main types of review question of relevance to NICE guidelines. These are listed in Table 3. 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 GC to formulate a recommendation.

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

Table 3: Best study design to answer each type of question

Type of question	Best primary study design
Effectiveness or other impact of an intervention	Randomised controlled trial (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 GC. 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

Update 2017

agreement between GC members (see Section 3.5.6) and the need for future research is specified.

3.5.13 The search process

3.5.1.14 Scoping searches

A broad preliminary search of the literature was undertaken in November 2014 to obtain an overview of the issues likely to be covered by the scope, and to help define key areas. The 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 H.

3.5.1.20 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. Searches were restricted to certain study designs if specified in the review protocol, and conducted in one or more of the following databases:

- CDSR, DARE
- CENTRAL
- Embase
- HTA database (technology assessments)
- MEDLINE/MEDLINE In-Process
- Psychological Information Database (PsycINFO)

With the exception of review questions 2.8 and 3.0, searches were undertaken in CENTRAL only with additional searching for pharmacological evidence published between 2004-2009 being carried out by the Cochrane Centre for Depression, Anxiety and Neurosis (CCDAN). For review questions 2.8 and 3.0, searches were undertaken in CDSR, DARE, CENTRAL, Embase, the HTA database, Medline and PsycINFO,

Where relevant 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 Guideline Committee to ensure that all possible relevant search terms were covered. In order to assure comprehensive coverage, search terms for depression 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 H.

3.5.1.35 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.

3.5.1.41 Search filters

To aid retrieval of relevant and sound studies, filters were used to limit searches to systematic reviews and RCTs. The search filters for systematic reviews and RCTs are adaptations of validated filters designed by the Health Information Research Unit (HIRU) at

Update 2017

McMaster University. Each filter comprises index terms relating to the study type(s) and associated text words for the methodological description of the design(s).

3.5.1.5 Date and language restrictions

Searches for systematic reviews and RCTs were undertaken for research published between January 2009 (the end of the search period for CG90) and April-May 2016. In addition, for pharmacological evidence, the CCDAN undertook searches of the literature for research published between 2004 and 2009 which was not updated in CG90. 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.

3.5.1.6 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) conducting searches in ClinicalTrials.gov for unpublished trial reports; (c) contacting included study authors for unpublished or incomplete datasets. 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 H.

3.5.1.7 Study selection and assessment of methodological quality

Titles and abstracts of studies identified by the searches were screened by two reviewers for inclusion against criteria, until a good inter-rater reliability had been observed (percentage agreement =>90% or Kappa statistics, K>0.60). Initially 10% of references were double-screened. If inter-rater agreement was good then the remaining references were screened by one reviewer. 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 a study database (standardised template created in Microsoft Excel). Eligible systematic reviews and RCTs were critically appraised for methodological quality (risk of bias) using the Cochrane risk of bias tool (in line with the *Developing NICE guidelines: the manual*).

3.5.1.8 Unpublished evidence

Stakeholders were invited to submit any relevant unpublished data using the call for evidence process set out in Developing NICE guidelines: the manual. Additionally, authors and principal investigators were approached for unpublished evidence. The GC 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 GC did not accept evidence submitted 'in confidence'. However, the GC 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.

Update 2017

3.5.2 Data extraction

3.5.2.1 Quantitative analysis

Study characteristics, aspects of methodological quality, and outcome data were extracted from all eligible studies, using Review Manager Version 5.3 (Cochrane Collaboration 2014) and an Excel-based form (see Appendix J).

In most circumstances, for any given outcome (continuous and dichotomous), where more than 50% of the number randomised to any group were missing or incomplete, the study was excluded from the analysis.

If some, but not all, of a study's participants were eligible for the review, for instance, mixed anxiety and depression diagnoses, and we were unable to obtain the appropriate disaggregated data, then we would include a study if at least 80% of its participants were eligible for the review.

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 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 GC was used to overcome difficulties with coding. At least 10% of data extraction was double-coded. Discrepancies or difficulties with coding were resolved through discussion between reviewers or the opinion of a third reviewer was sought.. Where consensus could not be reached, GC 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, Moore et al. 1996, Berlin 2001).

Update 2017

3.5.3 Evidence synthesis

The method used to synthesise evidence depended on the review question and availability and type of evidence (see Appendix F for full details).For questions about the effectiveness of interventions, network meta-analysis (NMA) or standard pairwise meta-analysis was used where appropriate, otherwise narrative methods were used with clinical advice from the GC. An overview of the NMA methodology used in this guideline is provided in Chapter 7; full details of NMA methods are described in Chapter 17 and Appendix N. In the absence of high-quality research, informal consensus processes were used (see Section 3.5.6).

3.5.4 Grading the quality of evidence

For questions about the effectiveness of interventions, the GRADE approach was used to grade the quality of evidence from group comparisons for each outcome (Guyatt, Oxman et al. 2011). The technical team produced GRADE evidence profiles (see below) using the GRADEpro guideline development tool, following advice set out in the GRADE handbook (Schünemann, Brożek et al. 2013). All staff doing GRADE ratings were trained, and calibration exercises were used to improve reliability (Mustafa, Santesso et al. 2013).

3.5.4.1 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 4 for an example of a completed 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 5 factors: limitations, inconsistency, indirectness, imprecision and publication bias. For the purposes of the guideline, each factor was evaluated using criteria provided in Table 5.

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 quality of the evidence for each outcome. Under the GRADE approach, the overall quality for each outcome is categorised into 1 of 4 groups (high, moderate, low, very low).

Update 2017

Table 4: Example of a GRADE evidence profile

Quality assessment							No. of patients		Effect			
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other consid-erations**	**Intervention**	**Control group**	**Relative (95% CI)**	**Absolute**	**Quality**	**Importance**
Outcome 1 (measured with: any valid method; better indicated by lower values)												
2	Randomised 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)	**moderate**	CRITICAL
Outcome 2 (measured with: any valid rating scale; better indicated by lower values)												
4	Randomised trials	Serious[2]	No serious inconsistency	No serious indirectness	Serious[1]	None	109	112	-	SMD 0.42 lower (0.69 to 0.16 lower)	**low**	CRITICAL
Outcome 3 (measured with: any valid rating scale; better indicated by lower values)												
26	Randomised trials	No serious risk of bias	Serious3	No serious indirectness	No serious imprecision	None	521/5597 (9.3%)	798/3339 (23.9%)	RR 0.43 (0.36 to 0.51)	136 fewer per 1000 (from 117 fewer to 153 fewer)	**moderate**	CRITICAL
Outcome 4 (measured with: any valid rating scale; better indicated by lower values)												

Update 2017

Quality assessment							No. of patients		Effect			
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other consid-erations**	**Intervention**	**Control group**	**Relative (95% CI)**	**Absolute**	**Quality**	**Importance**
5	Randomised 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

Notes:
[1] OIS (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.
CI = confidence interval; OIS = optimal information size; RR = risk ratio; SMD = standardised mean difference.

Update 2017

Table 5: 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 the Cochrane risk of bias tool (see Section 3.5.1).
Inconsistency	Unexplained heterogeneity of results.	Moderate or greater heterogeneity (using the methods suggested by GRADE[1])
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 GC 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 2 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 (a) no effect and (b) 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.

Notes:
[1] For heterogeneity, outcomes were downgraded once if $I^2 \geq 50\%$ and twice if $I^2 > 80\%$. If heterogeneity was found, subgroup analysis was performed using the pre-specified subgroups in the protocol (see Appendix F); if subgroup analysis did not explain the heterogeneity, a random-effects model was used and the outcome was downgraded.
GRADE = Grading of Recommendations Assessment, Development and Evaluation; OIS = optimal information size.

Update 2017

3.5.52 Presenting evidence to the Guideline Committee

Study characteristics tables and, where appropriate, forest plots generated with Review Manager Version 5.3 and GRADE summary of findings tables (see below) were presented to the GC.

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 GC. The range of effect estimates were included in the GRADE profile, and where appropriate, described narratively.

3.5.5.10 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 6). The tables provide anticipated comparative risks, which are especially useful when the baseline risk varies for different groups within the population.

Table 6: Example of a GRADE summary of findings table

Outcomes	No of Participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with intervention (95% CI)
Global impression: 1. no improvement – short term	102 (1 study)	**low**[1,2] due to risk of bias, imprecision	RR 0.89 (0.69 to 1.16)	725 per 1000	80 fewer per 1000 (from 225 fewer to 116 more)
Behaviour: 1. average change score Adaptive Behaviour Scale – medium term	101 (1 study)	**low**[1,2] due to risk of bias, imprecision		The mean behaviour score was 1	0.60 SDs lower (1 to 0.21 lower)
Adverse effects: 1. extrapyramidal symptoms – medium term	243 (2 studies)	**low**[1,2] due to risk of bias, imprecision	RR 0.34 (0.05 to 2.1)	33 per 1000	21 fewer per 1000 (from 31 fewer to 36 more)

Notes:
The basis for the assumed risk was the median control group risk across studies. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
[1] Generally unclear risk of bias and funded by manufacturer.
[2] OIS (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
CI = confidence interval; GRADE = Grading of Recommendations Assessment, Development and Evaluation; OIS = optimal information size; RR = risk ratio; SD = standard deviation.

Update 2017

3.5.6 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, an informal consensus process was adopted.

The process involved a group discussion of what is known about the issues. The views of the GC were synthesised narratively by a member of the review team, and circulated after the meeting. Feedback was used to revise the text, which was then included in the appropriate evidence review chapter.

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 covered in this 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

Developing NICE guidelines: the manual. Prioritisation of areas for economic modelling was a joint decision between the Health Economist and the GC. The rationale for prioritising review questions for economic modelling was set out in an economic plan agreed between NICE, the GC, 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 pharmacological, psychological, physical and combined interventions for adults with a new episode of less severe depression (RQ 2.1)
- cost effectiveness of pharmacological, psychological, physical and combined interventions for adults with a new episode of more severe depression (RQ 2.2)
- cost effectiveness of pharmacological, psychological and combined pharmacological and psychological interventions for preventing relapse in adults whose depression has responded to treatment (RQ 2.3)

In addition, literature on the health-related quality of life of people covered by this guideline was systematically searched to identify studies reporting appropriate utility scores 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

3.6.1.1 Scoping searches

A broad preliminary search of the literature was undertaken in November 2014 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.

3.6.1.2 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.

Update 2017

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 GC 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 interventions by authors in the titles and abstracts of records.

For standard mainstream bibliographic databases (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 I.

3.6.1.3 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.

3.6.1.4 Search filters

The search filter for health economics is an adaptation of a pre-tested strategy designed by the *Centre for Reviews and Dissemination* (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 I.

Update 2017

3.6.1.5 Date and language restrictions

Searches for economic evaluations and quality of life studies were undertaken for studies published between January 2002 and July 2016, with 2002 being used as a back date to capture pharmacological research not reviewed in CG90. After this point, studies were included only if they were judged by the GC 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.

3.6.1.6 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 I.

3.6.1.7 Inclusion criteria for economic studies

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

1. 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. For each review question and each strategy (intervention or service delivery model/setting), the focus of the economic literature review was on UK evidence.
    - For review questions that were supported by guideline economic modelling, only UK economic studies were included in the review.
    - For the remaining review questions that were not supported by economic modelling, UK evidence on each strategy was sought first; if no UK economic evidence was identified or the UK evidence was very thin (i.e. if it came from a single UK study or was characterised by very serious limitations), then a hierarchy of criteria were used to include studies in the economic review according to the country of origin, considering the similarities of each country's health system to the UK NHS, as follows:
        - Economic studies from Europe, Canada, Australia and New Zealand
        - Economic studies from the US
        - Economic studies from the remaining OECD countries (Chile, Mexico, Turkey, Israel, Japan, Korea)

    The described hierarchy for identification of eligible studies was agreed by the GC and the Health Economist and was followed until at least 2 economic studies were identified for each intervention or model of care considered in every review question; if less than 2 studies were identified, then studies meeting the next criterion in the hierarchy were sought.
2. Selection criteria based on types of clinical conditions and service users as well as interventions assessed were identical to the clinical literature review.
3. 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. Conference abstracts, poster presentations or dissertation abstracts were excluded.
4. Full economic evaluations that compared two or more relevant options and considered both costs and consequences were included in the review (i.e. (cost-utility, cost-effectiveness, cost-benefit or cost-consequence analyses)
5. Economic studies were included if they used clinical effectiveness data from a randomised or non-randomised clinical trial, a prospective cohort study, or a systematic review and meta-analysis of clinical studies. Economic analyses that utilised data from studies with a mirror-image design and studies that recruited participants retrospectively were not considered in the review, due to their lower methodological quality.
6. Studies that adopted a very narrow perspective, ignoring major categories of costs to the NHS, were excluded; for example studies that estimated exclusively intervention costs were considered non-informative to the guideline development process. In addition, studies that considered an employer's perspective and included only productivity losses and/or benefit payments were not included in the review.
7. Studies comparing healthcare costs of adults with depression receiving branded versus generic forms of drugs were not considered in the economic literature review.

3.6.1.8 Inclusion criteria for health state utility studies

1. Only studies from Organisation for Economic Co-operation and Development countries were included.
2. 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

Update 2017

that the study's data and results were extractable. Conference abstracts, poster presentations or dissertation abstracts were excluded.

3. To be included, studies should report utility data for specific health states associated with depression through the care pathway. Studies reporting an overall utility score for people with depression (and/or people without depression), who may have a mixture of depression-related health states or a range of symptom severity, were not considered.
4. HRQoL should be rated directly from adults with depression using the EQ-5D valued by the general UK population, according to NICE recommendations (NICE 2013: *Guide to the Methods of Technology Appraisal*). If no such studies were available, then a hierarchy of criteria were used to include studies in the review, as follows:
    - use of SF-6D utility data, derived using the UK algorithm for valuation (Brazier et al. 2002)
    - use of EQ-5D valued by a population of another country
    - use of another validated generic preference-based measure (PBM) [e.g. SF-6D valued by a non-UK population, HUI-3]
    - use of a condition-specific PBM valued by general population (UK data prioritised over non-UK ones) using time trade-off (TTO) or standard gamble (SG)
    - use of vignettes valued by the general population (UK data prioritised over non-UK ones) using TTO or SG
    - use of condition-specific PBM valued by service users (UK data prioritised over non-UK ones) using TTO or SG
    - use of vignettes valued by service users using TTO or SG, or direct service user valuations of their own HRQoL (UK data prioritised over non-UK ones).

Update 2017

3.6.24 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 2014: Developing NICE guidelines: the manual). 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. The completed methodology checklists for all economic evaluations that were included in the guideline are provided in Appendix P.

3.6.33 Presentation of economic evidence

Existing 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 Q. Methods and results of economic modelling undertaken alongside the guideline development process are provided in Chapter 13 and Chapter 14. 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 in Appendix R.

3.6.42 Results of the systematic search of economic literature

The titles of all studies identified by the systematic search of the literature (N=32,785) were screened for their relevance to the topic (that is, economic information and health state utility data relating to adults with depression). Three more studies were identified through the GC, which were unpublished at the time of the final search. References that were clearly not relevant were excluded first. The abstracts of all potentially relevant studies (635 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 1 study, or had been updated in more recent publications were subsequently excluded; studies not meeting the inclusion criteria for hierarchy of settings/countries were subsequently excluded. Economic evaluations eligible for inclusion (47 cost effectiveness studies in 53 publications, of which 2 included utility data as well, and another 4 studies providing utility data) were then appraised for their applicability and quality using the methodology checklist for economic evaluations. Finally, those studies that fully or partially met the applicability and quality criteria set by NICE were considered at formulation of the guideline recommendations. The flowchart of the studies considered in the systematic review of the economic literature is shown in Appendix O. The list of excluded studies after obtaining full text or following the hierarchy of countries/settings is provided in Appendix S.

3.7 From evidence to recommendations

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

Finally, to show clearly how the GC moved from the evidence to the recommendations, each chapter (or sub-section) has a section called 'recommendations and link to evidence'. Underpinning this section is the concept of the 'strength' of a recommendation (Schünemann, Best et al. 2003). This takes into account the quality of the evidence but is conceptually different. Some recommendations are 'strong' in that the GC 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 GC 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.

Where the GC 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 G.

Update 2017

3.8 Methods for reviewing experience of care

3.8.1 Introduction

The chapter on experience of care (Chapter 4) presents three different types of evidence: personal accounts that were collected by the service user and carer members of the GDG; interviews from the Healthtalkonline website (www.healthtalkonline.org); and review of the qualitative literature.

3.8.2 Personal accounts

The authors of the personal accounts were contacted primarily through the service user and carer representatives on the GDG, and through various agencies with access to people with depression. In approaching these individuals, the GDG attempted to assemble a range of individual experience that reflected what the GDG considered to be important aspects of the care and treatment of people with depression. All individuals who were approached to write the accounts were asked to consider a number of questions (see Chapter 4) prepared by a service user and carer topic group4 which oversaw this aspect of the guideline work. Each individual signed a consent form giving permission for their account to be reproduced in this guideline. All personal accounts were read by the members of the service user and carer topic group, and the review team; if necessary, the authors of the accounts were contacted again if parts of their account were unclear or ambiguous, or where it was thought that further information would be helpful. Any changes made for clarity were approved by the authors of the accounts. The full text of the accounts is reproduced in this guideline. The personal accounts were read again by the service user and carer topic group, and the review team, and themes were identified. These themes were developed and reviewed by the topic group and then incorporated in a combined summary with the evidence from the other two sources below.

3.8.3 Interviews from Healthtalkonline

Using the interviews of people with depression available from healthtalkonline.org, the review team analysed the available data and identified emergent themes. Each transcript was read and re-read, and sections of the text were collected under different headings using a qualitative software program (NVivo). Two reviewers independently coded the data and all themes were discussed to generate a list of the main themes. The evidence is presented in the form of these themes, with selected quotations from the interviews. The methods used to synthesise the qualitative data are in line with good practice (Braun & Clarke 2006).

3.8.4 Review of the qualitative literature

A systematic search for published reviews of relevant qualitative studies of people with depression was undertaken using standard NCCMH procedures as described in the other evidence chapters. Reviews were sought of qualitative studies that used relevant first-hand experiences of people with depression and their families or carers. The GDG did not specify a particular outcome. Instead, the review was concerned with any narrative data that highlighted the experience of care. The evidence is presented in the form of themes, which were again developed and reviewed by the topic group.

3.8.5 From evidence to recommendations

The themes emerging from the personal accounts, the qualitative analysis of the Healthtalkonline transcripts and the literature review were reviewed by the topic group. They are summarised in Chapter 4 and this summary provides the evidence for the recommendations that appear in that chapter.

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

Update 2017

- 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, so a 'national' organisation is defined as 1 that represents England, or has a commercial interest in England.

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
- commenting on the draft of the guideline.

3.107 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 D) 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 GC finalised the recommendations and the NGA produced the final documents. These were then submitted to NICE for a final check. Any errors were corrected by the NGA, then the guideline was formally approved by NICE and issued as guidance to the NHS in England.

Update 2017

4 Experience of care

4.1 Introduction

This chapter provides an overview of the experience of people with depression and their families/carers. In the first two sections are first-hand personal accounts written by people with depression and carers, which provide some experiences of having the diagnosis, accessing services, having treatment and caring for someone with depression. It should be noted that these accounts are not representative of the experiences of people with depression and therefore can only ever be illustrative. This is followed by a qualitative analysis of transcripts of people with depression from the Healthtalkonline website (http://www.healthtalk.org/) and a review of the qualitative literature of the experience of people with depression. There is then a summary of the themes emerging from the personal accounts, the Healthtalkonline transcripts and the literature review, which provides a basis for the recommendations, which appear in the final section.

4.2 Personal accounts – people with depression

4.2.1 Introduction

The writers of the personal accounts were contacted primarily through the service user and carer representatives on the GDG and through various agencies that had access to people with depression. The people who were approached to write the accounts were asked to consider a number of questions when composing their narratives. These included:

- When were you diagnosed with depression and how old were you?
- How did you feel about the diagnosis? How has your diagnosis affected you in terms of stigma and within your community?
- Do you think that any life experiences led to the onset of the condition? If so, please describe if you feel able to do so.
- When did you seek help from the NHS and whom did you contact? (Please describe this first contact.) What helped or did not help you gain access to services? If you did not personally seek help, please explain how you gained access to services.
- What possible treatments were discussed with you?
- Do you have any language support needs, including needing help with reading or speaking English? If so, did this have an impact on your receiving or understanding a diagnosis of depression or receiving treatment?
- What treatment(s) did you receive? Please describe both drug treatment and psychological therapy.
- Was the treatment(s) helpful? (Please describe what worked for you and what didn't work for you.)
- How would you describe your relationship with your practitioner(s)? (GP/community psychiatric nurse/psychiatrist, and so on.)
- Did you use any other approaches to help your depression in addition to those provided by NHS services, for example private treatment? If so please describe what was helpful and not helpful.
- Did you attend a support group and was this helpful? Did any people close to you help and support you?
- How has the nature of the condition changed over time?
- How do you feel now?

- If your condition has improved, do you use any strategies to help you to stay well? If so, please describe these strategies.
- In what ways has depression affected your everyday life (such as schooling, employment and making relationships) and the lives of those close to you?

Each author signed a consent form allowing the account to be reproduced in this guideline. Seven personal accounts from people with depression were received in total. Although the questions were aimed at people with any form of depression, all of the personal accounts received were from people who have/have had severe and chronic depression, spanning many years. The themes that are most frequently expressed in the testimonies include trauma or conflict in childhood as a perceived cause of depression; the need for long-term psychotherapy for people with severe and chronic depression; the need to take personal responsibility for and understand the illness to improve outcomes; issues around diversity; paid and unpaid employment as an important part of the recovery process; the negative impact on daily functioning; concerns regarding stigma and discrimination in the workplace; and the relationship between people with depression and professionals.

4.2.26 Personal account A

I was 23 when I was first diagnosed with depression, 35 when diagnosed with major depressive disorder and 43 when diagnosed with dysthymia. However, my first experience of suffering with depression was most probably as a teenager, living in a chaotic household with a parent with alcoholism and a narcissistic personality disorder.

The first treatment I had was when I was 23 with a wonderful GP who told me he had had depression and a breakdown at medical school. He enabled me to go to see him whenever I wanted, to talk to him for 10 to 15 minutes every week. I was also on an antidepressant and tranquilliser for instant tranquillisation whenever I felt miserable. The depression passed within 4 to 5 months. I always think of the GP fondly as a life saver.

For the next few years I used therapy to deal with my depression, low self-esteem and my underlying childhood issues, each year becoming more confident. During my childhood I had had to deal constantly with my mother's tempers, mood swings and cruelty, so I had to learn in therapy how to deal with my own emotions from scratch. Initially I had 3 years of gestalt therapy with a wonderful therapist who came recommended by a friend. I then had psychodynamic psychotherapy for 4 years (while I also ran a self-help group for women). I found this psychotherapist from the UKCP list. During this period I also worked with teenagers and I found hard work to be a great help in having something to focus on and enhance my self-esteem.

In my 30s, however, I had a major depressive episode and I booked myself into hospital which I now see as a big mistake as it was not therapeutic by any means, but my understanding of what hospital offered was not known to me. I had been having some housing problems, family life was difficult and I had been working very long hours at work to solve all of these problems. I knew that I was at danger point. I was given antidepressants, an antipsychotic, a mood stabiliser and benzodiazepines. I was offered no therapeutic help and I found the system of nursing within the ward very damaging – they just observed the patients and didn't talk to us. So I was just left with my depressed thoughts for 11 weeks. I came out and went back to work.

I also didn't realise that there was stigma around these matters, and I had been open with my friends about being depressed and in hospital. Overnight I lost two thirds of my friends and social contacts. This left me feeling very distressed, ashamed and humiliated. Also, within my family, my illness was exploited by my still-crazy mother, to undermine and separate me from any compassion I could expect. This has changed gradually over the years, but it took a long time to heal.

At work, although I was employed in the care environment, some people were not keen about me returning to work. I was marginalised from external meetings for quite some time and my role was circumscribed. This changed over time, but I don't think I should have had to 're-prove' myself as if I had been in prison. But I kept quiet and got on with it. I learnt that it's best to hide having depression, to avoid the stigma. Subsequently, I have discovered through my own experience and working with service users, that it's still best to hide having depression (or indeed any other mental illness) if you want to get a job and keep it.

I have had two recurrences of major depressive disorder. I had to give up work in 1998 to battle with it full time for a couple of years. I begged to have psychotherapy but I now couldn't afford to pay for it myself. I was tried on a series of drugs over a 7-year period: six different antidepressants and various mood stabilisers, tranquillisers, and so on. I got a job in 2000, but I could barely hold a conversation I was so drugged up. It was sheer force of will that got me up and out each day. I was swimming and eventually was able to pay for my own psychotherapy, and gradually the major depression I had been in for 4 to 5 years lifted in 2002. Throughout this time I had battled with pervasive suicidal feelings and only my personal strength got me through. Just getting off the huge amounts of medication was a feat I am proud of in itself, in addition to overcoming the depression caused by childhood issues and living a normal positive life which the medication, not to mention the illness, nearly took from me completely.

I also had a wonderful GP in 2002 to 2003, who took it upon himself to (in his words) 'have a go at' at my consultant psychiatrist for half an hour on the phone about the cocktail of drugs I was taking. Being on a level of medication that was unnecessary and toxic, I had put on seven and a half stone since 2005 and I was threatened with high blood pressure and impaired glucose syndrome. My GP helped me get off this cocktail of unnecessary medication.

Not being drugged up freed me and enabled me to function at work, as I had previously done, and it 'woke' me up. The threatened 'relapse' has never happened. My self-esteem issues over my depression and weight had left me anxious though, and after an 18-month battle involving Mind and my psychiatrist, I got cognitive behavioural therapy (CBT) in 2004. This was even more wonderful in aiding my recovery and I had one session per week for a year working on my anxiety phobias. The psychologist was a wonderful professional who had faith in me and together we worked very hard overcoming the deep beliefs that I had held and which prevented me leading a full, well life.

I have been having psychotherapy again since 2005, working on the final bits of damage done to me by my alcoholic, narcissistic mother. It is hard work but my personal stamina increases all the time. This therapy would not be available in the local mental health trust – there is only one course of psychotherapy available (1 year per patient). Even with lifelong illness you get one 'go' at it. Where I currently live, patients cannot choose whether they would prefer a male or female therapist, nor the style of training they would want their therapist to have had. Choosing a therapist is as important as choosing a GP. Within the NHS there is still a culture that if you don't take any therapist, you are treatment resistant. I have always preferred a woman therapist, and one psychodynamically or psychoanalytically trained.

My psychotherapist is helping me with positive attachment and parenting techniques to get to the point I should have been at, and forming a positive attachment in the psychotherapeutic environment. This enables me to build confidence and be the person I should be, making the most of my abilities and relationships in the present. I am also learning self-analysis and skills building to enable me to keep an eye on stresses and challenges, to self-manage and keep well.

My psychiatrist, who I had from 1995 to 2005, now agrees with me that psychotherapy, building my career and not being on any drugs, have been the best for me in my recovery. She is of the 'old school' and took a lot of convincing, but at some point, she turned her ideas

around about me and what I was able to achieve. She still confirms I was very ill, but that with my hard work I have completely changed my life around and, in her terms, I am unlikely to relapse. My psychiatrist put this in writing to my GP in 2006.

Stigma remains a problem however. It is worse if the negative attitudes are expressed by GPs and other medical practitioners. Even now assumptions seem to be made when I have outpatient appointments for physical ailments because computerisation of records has meant even though I have recovered, major depressive disorder is on my records everywhere. I can sometimes see a doctor's face drop when they get to that point – some are not very good at hiding it. In 2006 I was turned away from a gastro clinic and told that my stomach pain and weight loss were because of depression and that the NHS couldn't help me. I complained and the resulting CT scan showed I had cancer which when removed 6 weeks later was at stage 2. I feel quite sick thinking of how many people with depression and mental illness, especially those who are less articulate and bolshie than me, could be being turned away because of the lack of understanding. If I had listened to that doctor in 2006, I would be dead now – and all because I have had depression, not for any other reason.

4.2.36 Personal account B

I first consulted my original GP in the spring of 2006, when I was 55, because of symptoms of what I felt was very severe and prolonged depression. I had experienced a rapid series of distressing life events (a complex bereavement leading to feelings of alienation and isolation) and I had no support. I was working freelance as a trainer but no longer able to seek work and so I was without an income.

I had already tried to help myself for 6 months and had bought many so-called self-help books. I have a Master's degree in social work and at one time taught counselling skills. I am familiar with rational emotive therapy, CBT, person-centred therapy, transactional analysis, and so on. I understand the efficacy of exercise, diet, positive thinking and relaxation. The major problem is that one cannot actually do these things when depressed and I believe those who have not been depressed cannot truly comprehend this at all. I am also conscious that any so-called emotional problems affect the way one is perceived and addressed. Because of this, I was very reluctant indeed to seek help and many of my fears were in fact confirmed.

The GP whom I first saw spent more time looking at his computer than me. He asked 'are you depressed?' I told him I was sufficiently distressed to consult a GP. Having said he could refer me to the mental health team, he said that they were 'not very good' and gave me a card for a private counsellor. He told me to complete a 'HADS' test in the waiting room and put it under his door. He offered no medication and no follow-up appointment. I sat in my car in the car park crying for 2 hours before I could drive home.

However, I made an appointment with the private counsellor, although I was anxious about the cost. But I felt I had to try and help myself. The counsellor was a very nice woman but I felt I was not being assessed. She talked a great deal about her upcoming wedding and for half a session explained the essentials of transactional analysis (which I've taught). I also felt that conclusions were drawn rapidly and inaccurately. She told me to keep a diary of angry feelings and never referred to it again. She explained that 'if you haven't had an adolescent rebellion you have one in middle age' and told me to 'get rid of' people who were draining me. This is not entirely bad advice but much too crude. I got the impression she was talking about her own life, not mine. I felt very much more unsettled at the end of each session than when I had arrived.

After three sessions I found another counsellor, who was better than the first but I could not afford to continue the sessions or to travel to see him. Again I found that the counsellor seemed to have a favourite model of human behaviour. I was later even more annoyed when the difficulties with the counsellors were explained away by a mental health team worker as a

disturbance of mine in facing the issues. I felt much worse afterwards knowing this and that I could not improve the situation.

Eventually I began a method of self-counselling: occasionally speaking aloud to myself in a deliberate effort to calm myself down since I knew that depression can be a result of over-stimulation.

Fortunately, in the summer of 2006, I was able to change my GP. The new GP provided much more help but unfortunately the initial medication (citalopram), which I took for 4 months, made no difference to me at all.

My new GP referred me again for counselling at the surgery. There was a waiting list: I attended the first session and then there was a gap of some weeks (which was at the end of 2006). I found it disturbing to have to talk to a stranger yet again. The sessions often ended with an emotionally laden question or the advice given was more appropriate for a much older bereaved person. I did very little talking and I could not summon the energy to constantly correct the assumptions being made which, again, seemed based on the counsellor's own life. I attended just a few sessions and then decided that this was a waste of resources.

I felt that if someone would just skilfully listen and question (as I thought good counselling did) I could sort things out myself. My own reasonably sound knowledge of counselling actually seemed to be a disadvantage to me and I had to learn to keep quiet. I still needed help, had very little external support, and my GP was offering what was available so I felt I had to accept it, but it was not even close to what I needed.

In February 2007 I got into a very distressed state but could not get an appointment with any GP although I phoned the surgery four times. The one friend who knows about my condition then took me to the surgery. I now know that I was quite seriously ill at this point. But one can only go to the surgery when one feels capable of doing so. Appointments had to be made on the day at 8.30 a.m. which was one of the worst times for me. So then appointments had to be made a few days ahead. One needs to be able to access help when one needs it during the bad times. In the end it was a registrar GP who saw me in this deeply distressed state. Even then I felt guilty for someone seeing me 'as an emergency' and I felt very bad about that. He was, however, quite good and he referred me again to the mental health team.

The registrar changed my medication to escitalopram. I was deeply grateful as my GP had kept telling me to continue the citalopram and wait for it to take effect. The escitalopram was beneficial and I have continued with it for over a year. I still seem to need this medication. I feel that getting the medication right and promptly at the virulent stage of the depression is vital. I also feel that I was quite poorly and was left to 'wait' to see if I would get better.

Prior to my mental health team assessment interview in May 2007 (the GP registrar I saw in February had written again to the team to ask for an early appointment) I was in a very foggy state and was particularly vulnerable. However, I think that I expressed the issues quite clearly in the limited time. The interviewer described himself as a nurse, said he was trying to clarify why I was there and at one point told me I looked 'alright', which was frustratingly puzzling to me and based on no knowledge of me whatsoever. I quickly lost confidence in my interviewer. He said, 'Yes, I've had bereavements too' and 'I don't know why you have been referred', which was very unhelpful. He also told me I had to 'negotiate' if the counselling is not right. How can someone who is seriously depressed negotiate?

I was also given the Aaron Beck tick box-type diagnostic tool which I found confusing. (For example 'loss of appetite' is difficult to answer; a lot of people who are depressed have 'abnormal appetite'.) I find these tools very simplistic.

I left this appointment and began crying immediately – again I could not drive home for an hour. I took extra medication to try and cope. I called the mental health team and was told that I was bound to get upset 'as I was talking about upsetting things'. Again, the problem is

presented as being because of the vulnerability of the patient rather than the competence of the interviewer.

My GP had said that she would be able to refer me to a psychologist but that first I had to be referred to the mental health team. I found this very disappointing and also embarrassing. I was going to have to tell yet another person about my life. When after many weeks I got to see the mental health team counsellor in June 2007 she told me the sessions were for 6 weeks so I knew immediately I could not be helped in this short time: I was taught 'relaxation training' which was inadequate for my needs. It was like offering aspirin for appendicitis. I had to miss one of the six sessions because I was not well enough to attend.

With every other (physical) condition for which I have been referred I have been seen by a consultant at least once. But with a mental health problem, which was the one life-threatening condition which I had, I was referred by a GP and seen by a nurse (who thought I 'looked ok'). This meant that I had problems getting my pension (money problems started to become a major factor when my savings diminished). The occupational health professional said I had to have a consultant diagnosis; but it was almost a year before I could see a psychiatrist for a formal diagnosis, which my former employer paid for.

I at last saw a consultant psychiatrist privately in January 2008. She diagnosed me with post-traumatic stress (I had been severely bullied at work before I left 10 years ago) leading to severe depression. While perhaps dismal, it was a relief to have the diagnosis and it does validate my experience. The psychiatrist saw me for two sessions but explained that she could not see me again (as this was, I expect, very expensive). She did provide details of a freelance psychologist, but told me that I would have to see her privately. I saw this psychologist twice paying £75 each session but just could not afford any further sessions. I have had no further treatment other than the medication. As my GP said very recently, there is no other help available, just 'short fix' stuff.

Over the past 2 years I have had to share my personal details over and over again with about 12 strangers, half of them doctors 'assessing' me. My GP has done her best, but has only so much time, and one wants to be a 'good' patient. At one point I stopped driving as I knew that I was not safe to do so. I told my GP about this but she said I would feel a sense of achievement if I continued to drive! This greatly concerned me. Also, I felt no 'sense of achievement': a lack of achievement is not one of my problems. I felt that my self-report was not being taken seriously and I was very confused about how I could present myself to make myself understood.

I was never clear about the role of the mental health team or what the 'variety of options on offer' actually was (in fact other than counselling there was 'nothing else available'). It was not recognised that I was in a deep fog, akin to being in another universe, and was finding it very hard to concentrate on what was being said. The more contacts I had, the more distressed I felt.

Up until 6 or 7 months ago I was feeling as if in a parallel universe, and at one point as if I was living under water. I could not 'wake up' from dreams, and very unusually for me I could not get up until 10 am on some days. I felt profound grief.

I now have far less faith in getting help so I do not know what I would do if things become worse. I was helped by seeing the consultant psychiatrist and I felt much better having been taken seriously. One problem was being not being able to work.

My own coping strategies are mainly avoiding known triggers, self-monitoring and trying to get proper nutrition. I also swim every day. Distraction helps if I can stop the circularity of thoughts. My everyday life is affected as I am much less outgoing now. I have been 'let down' so many times that I do not want to make the approach now. I am mostly happier on my own though I am also gregarious and socially skilled. I feel a little embarrassed that I do not have the things other people of my acquaintance have (family relationships and so on)

and so I cannot talk the currency of that group (children and grandchildren). But I am more accepting of my own isolation/difference from other people. However, I do fear being destabilised by even small life events in the future as I know I am vulnerable and don't manage such challenges well.

4.2.45 Personal account C

Life experiences have definitely led to the onset of depression. I had an accident as a child which affected my eyesight and I have been visually impaired all my teenage and adult life. After I lost my sight I felt I was rejected as a child and teenager by my family, which was exacerbated by being sent away from home to be educated at a school for blind people. As the eldest of four children I bore the brunt of my father's aggression and when I was older had to work in the family business for long hours and was punished at whim.

Because of my impaired sight I have had problems with sensitive hearing that made my life hell. I felt like a prisoner and as if I was being tortured by everybody and everything with so much noise around me.

I was admitted to a psychiatric unit at the age of 30 because I was suicidal. This was due to a variety of reasons which had been building up to that time. The main complication was that my wife was expecting a baby and we were not getting on and constantly arguing. I felt totally lost, I had no friends and there was no support for my depression. Because of my past experience I couldn't go to my parents or brother or sisters who lived near me. I felt totally isolated and not wanted by anybody. Although I received a diagnosis of depression this was not fully explained to me and it didn't do any good because ultimately the staff weren't equipped to help me or my family. They couldn't give proper information in a manner that my family could accept or understand, or communicate with them effectively, and there has been no support since then. I spent 6 days there and was medicated. The treatment was ultimately not helpful because there was no follow-up support.

In 1992 I attended a college for the blind for training in the hope that I would be able to get a job. Unfortunately this didn't happen because I was so unprepared, was having emotional breakdowns, and had too much to cope with at college. I was sent to a local hospital by a doctor from the college and was diagnosed with problematic depression and was given more practical help than previously: I had some psychotherapy, relaxation classes and exercise for my neck. At the end of the college year I was advised to take a break of a few months. This was a very hard time and a struggle for me – both the college and the job centre rejected me by saying they couldn't help me until I was stable.

There is a definite stigma towards mental health problems in my community, which is Muslim. Nobody seemed to want to understand about my diagnosis and I didn't feel I could talk to anybody because people are not equipped to provide support. They believe in leaving it to the power of prayer. When I approached an Imam in a local mosque about a personal problem within the family I was told that religion would resolve it. He stirred up more trouble by visiting the family member with whom I was having difficulties.

I have felt like an outsider and have suffered rejection after rejection. I have been rejected from services, society and family. I feel like my life is messed up physically, mentally, socially and financially, and in terms of work and education.

I had a severe breakdown last year and am concerned about relapse and was referred twice by my GP to the community mental health team. I was not seen by them. I feel like I am wasting my time trying. I feel like I am being pushed back. I am in a situation where I need the support of a therapeutic community or at the very least a safe place where I am able to get away from family pressures.

My relationship with my current GP is better at the moment. I don't have regular check-ups or practical support but I get help with medication and an occasional chat if I bring the subject

up. My GP was a bit more helpful when I had my breakdown. The CMHT did not do a good job of giving practical help: instead I was passed on to voluntary groups who were not fully equipped to offer support in a crisis or if I need help for referral from my GP to the CMHT again. It feels like a vicious circle: I have had a total of five breakdowns and have attempted suicide. But this seems to mean nothing to them. The only psychiatrist I have ever met told me that I would have to sort my problems out for myself. He literally let me wander the streets. I felt so bad I could have jumped off the roof. But perhaps God saved me.

I have therefore spent the last 15 years working on complementary therapies and any improvement in my condition is due to the work that I have done. It is more to do with faith and spirituality rather than religion. I feel closer to God now and feel protected. Many times I wanted to die and take the jump and I was saved. So I think I am meant to live and survive – there is a purpose for me otherwise I would have given up long ago or gone to prison or got on drugs and alcohol. So I thank God I have not gone down those roads.

The self-help techniques I have used have included positive affirmation, relaxation and emotional freedom therapy. I have also received qualifications in holistic therapies. I have been instrumental in setting up a local mental health drop-in centre and I am also a director of a local division of Mind and am standing as the BME representative on Mind Link. (I was able to access some CBT through Mind.) I have joined different groups, for example, a bowls club for blind people, and I have friends who have provided me with support.

But despite all this activity I am still disillusioned by the attitude of organisations that are meant to be dealing with mental health problems. I have a lot to offer despite no help being offered to me.

My feelings of alienation and isolation are exacerbated by family members who appear to have little appreciation of how difficult life is for me. I feel very isolated because my sensitive hearing makes me nervous and anxious in public places.

Depression has infected every part of my life. It has slowed me down, led to loss of self-esteem and made it difficult for me to get work.

4.2.58 Personal account D

The depression started when I was young (I am now 57). I came from a poor background – my father was diagnosed with bipolar disorder when he was in the army during the Second World War and after being discharged he spent a year in a psychiatric hospital. He couldn't work most of the time. My father also suffered from agoraphobia, so I ran errands for him – I was his 'skivvy'. My father had bad mood swings, which affected my mother, my siblings and me. He never gave any praise, and he never once said that he loved me or my mother. I missed school in order to care for him or because he had hit me so hard I had a black eye and couldn't go to school. I found it hard to learn at school and later I found out that I had dyslexia.

When I started puberty I felt different from other people. I felt as though I was not as good as the next person, which stemmed from my upbringing. There were a lot of kids at school living in poverty but life with my father made me feel very inadequate. When I was 15 or 16 years old my father tried to kill my mother when he found out she was having a relationship with another man. I felt as if I was always protecting my mother from my father. Both my siblings, who are older than me, married young to get away from my father.

I knew my feelings were different from those of other people so I went to see the doctor by myself when I was 16. The doctor knew immediately that I was suffering from depression. Because of my low self-esteem I couldn't hold a job down because I felt as if I was not good enough to do anything. I was constantly comparing myself to other people. I felt at the time that life wasn't worth living – I thought that practically it would be better to throw myself under a bus. If I hadn't gone to the doctor I would have killed myself. It was a relief to know that my

depression could be understood, if not treated, and to speak to someone who knew what I was talking about.

I was first prescribed diazepam, which made me feel good because I was out of it. I was prescribed one tablet a day but I took three or four. I couldn't work but at least it was a lift and that is what I felt I needed. I was on diazepam for about 6 to 9 months and then I came off it. I tried to look for a job but my feelings of inadequacy and paranoia returned: I felt as if people were looking at me and talking about me. I found it difficult to go outside and became agoraphobic.

Nothing else was offered to treat me so I treated myself by using cannabis, speed and barbiturates. Eventually I found a job I liked and when I was 18 years old I started having serious relationships. I was still living at home then and stayed to protect my mother as my father was still beating her, and I didn't want to take anyone home as I was ashamed of my father.

I finally left home at age 21 when I got married; I felt as if life was taking off. I was happily married and away from my father and it felt like depression was behind me. I loved my wife and that was enough in life. Children completed the marriage. By the time I was in my early 30s I was working in the building trade as a site manager and I was earning good money for the first time. I was determined not to be like my father and I appreciated what I had. I felt that there was a crater in my life where my father should have been. I didn't have anyone to look up to – no one to build a personality around. My personality only grew when I got married.

My Dad died in 1983. I stood by his grave and I couldn't cry. I battered myself with questions: what is the matter with me? I was consumed with all the thoughts of what had happened in the past. I felt numb about it all; it seemed like there was a massive void. I felt like I had never had a Dad and I became very good friends with a man in his 60s who I tried to adopt as a father.

In the following year my wife was diagnosed with schizophrenia. She was 28 at the time. My wife's illness made me feel depressed but I couldn't show it. I felt as though I had lost my wife and there was just a shell of a person there who used to be my wife. The illness was like a bereavement. I was offered antidepressants but I didn't take them as I didn't want my wife to see them. I was trying to keep it together but she believed I was having a nervous breakdown. Throughout her illness I was on an adrenaline rush. I was working flat out and didn't have time to think about myself. I was a machine trying to keep my family together: looking after my wife and kids and working. In the end I took time off work. I needed some emotional help and I needed someone to talk to. There was no time for myself and I stopped communicating with people.

After my wife had sufficiently recovered from her first episode of schizophrenia (it took about 9 or 10 months), I realised how badly it had affected me. I thought about what it had taken out of me and I would sink into depression and phone up the Samaritans. I went to see my GP a few times during this time and they were sympathetic to what I was going through. I started taking amitriptyline and I also saw a counsellor for 3 months. The counsellor was better than the antidepressants. It gave me a good lift. This lasted for a few months before I began to feel low again. For a few years I was in a cycle of relapsing and recovering – I was up and down like a yo-yo. I couldn't set a course for a life; everything had been completely obliterated by illness.

But my wife was feeling better and we wanted more children so the doctors took her off her depot antipsychotics and antidepressants. When she became pregnant she was happy and like she used to be before the illness. In 1987 my youngest son was born but 4 months after his birth my wife became very ill; she was hearing voices and it was as if the gates of hell were opened and everything came out. She was hospitalised and I stopped working and looked after the baby – it was like being a one-parent family.

Shortly after this I was diagnosed with asthma, which was considered by my doctors to be my major illness rather than depression. The asthma hit me hard as I was my wife's carer and I looked after the children. I also began to have panic attacks. Although I was convincing my wife that I was coping, this was just a mask. I felt as if I had become invisible, that my purpose was to make someone else become well. I did not see that there was something wrong with me. Then one day I was pushing a trolley around the supermarket and I thought 'I don't want to die in a supermarket; I don't want to die in between the bleach and the biscuits.' This happened several times around this period. I didn't go to doctors as I thought they would think I was nuts.

In 1997 my wife relapsed again and it affected our youngest son very badly as he had not seen his mother this way before. He was badly bullied at school for having a mother who was a 'nutter' and got very depressed. When he was 15 (in 2003) our son was also diagnosed with schizophrenia. I got depressed about what was happening to my son because I didn't want him to go through the same things that his mother and I had been through.

Although people think that I am stable, I recognise that I will never be free of depression but as I get older I understand more about it. I don't want to kill myself. I care for both my son and my wife and I will never turn away from them. I become more depressed when there is a crisis – and there always seems to be a crisis in my family. But I have accepted my depression as I have lived with it for so long; it's like an old nemesis. It's a part of me.

Eighteen months ago I was taking venlafaxine but I am not currently been treated for depression. To be honest, I hate taking tablets. When I was first ill I thought I was a lunatic because I was taking tablets. If I do need help I find that counselling is best for me, although I have not seen a therapist for a few years. I can now recognise when I am becoming depressed. It's a waiting game. I get black days when I wake up in the morning and I am totally unmotivated and I couldn't even care if I won the lottery – it would make no difference because I feel so lousy. If I feel like this for more than one day then I start to worry and I know I am depressed. To try and cope with the symptoms I grin and bear it or I try doing something different – getting away from mundane routine.

I am now able to talk to my wife about being depressed rather than trying to hide it from her and I talk to lots of other depressed people, which, for me, is like a form of counselling. I got involved with voluntary groups when my wife got schizophrenia: I am the chair of one voluntary organisation and I work for another, and I do a lot of media work. The horrid feeling of not being as good as other people is not there now because I feel that I am helping.

I am particularly interested in the political side of how people with mental health problems are treated. I believe that my depression was caused by my childhood experiences, but depression is such an individual illness – it has got many different faces and it can be caused by many different things. Therefore should people with depression be treated in the same way? I am encouraged to see that a lot of resources are being put into providing CBT for people with depression, but CBT is not the right treatment for everyone with depression and this needs to be recognised.

4.2.61 Personal account E

I was 27 years old when I was first diagnosed with depression, 14 years ago. I think I started to get depressed 6 years prior to diagnosis, I just didn't know it at the time.

At first, I was relieved at the diagnosis. I had gone to the doctors knowing something was wrong, but not knowing what it was. I was offered counselling and/or medication. I knew that I had to have medication, as it would make me feel better more quickly. I had already withdrawn from my friends and community (due to the depression) so in terms of stigma, there was none, though I didn't tell family, because they wouldn't have understood.

I knew that this ‘breakdown’ occurred due to the events that had happened the previous 18 months: the sudden deaths of two close friends and my grandmother, being made redundant from my part-time job, ending a 6-year relationship with my boyfriend, and then being physically assaulted.

Without doubt, my childhood experiences have also contributed to a life of depression. My mother died when I was 5 and after that my two younger brothers and I were not allowed to talk about her. My Dad remarried a woman with three children, but it was not long before my Dad and stepmother hated each other, and were physically and emotionally cruel to each other. My Dad hated her children, and was physically and emotionally cruel to them, and my stepmother hated my brothers and me, and was physically and emotionally cruel to us. One of my stepsisters sexually abused my youngest brother and me.

A month or so after starting medication, I did not feel any better, so was given counselling immediately. I established a good and trusting relationship with the counsellor who helped me to understand what was happening to me. However, I plummeted further, and was seen by a psychiatrist who allocated me a CPN, who I saw for around 18 months, until I was able to slowly start rebuilding my life. When my ‘time’ was up seeing the counsellor, I saw a psychologist for the following 18 months. I was also prescribed an antipsychotic drug, but I felt like a zombie and could not look after my daughter, so did not take it often.

Of the professionals listed above, without doubt the CPN helped the most; I had a good relationship with her. When I was at my most depressed, I was seeing the psychologist, but I was in no fit state to engage in any meaningful therapy, as I was too ill.

As well as the treatments listed above, while I was having counselling I was told that I should attend a women’s group, run by my counsellor through the NHS. I attended and it helped much more than I realised at the time in that I formed friendships that were very supportive. However, in terms of therapeutic input it did nothing – people would talk about their week and how awful life was, but I couldn’t do that. How could I tell people that I had spent the week trying not to kill myself, when that was all I wanted to do? It was not that I wanted to die, but I could see no other way of stopping the pain. Depression filled every second of every minute of every day, and it was unbearable. I was fortunate in that I was able to sleep a lot (up to 15 hours a day), though time still went slowly. Reading books about depression and self-help gave me an understanding of what was happening to me.

On one occasion I went to a voluntary agency support group, but I couldn’t accept at that time that depression would be part of my life forever: I found it difficult to listen to others about how they were managing their lives living with depression. I thought I was going to get better and it would never come back again – how naïve was I?

Over the years, I have been prescribed most of the SSRIs. They worked to varying degrees, but the most distressing aspect for me is that they all seem to affect my memory and articulation. I have learnt to live with this, but am aware of the limitations this poses for me, especially at work. I did receive further counselling on one occasion, by the NHS, but it was not particularly helpful, as it did not get to the root of the depression.

Over the last 2 years I have paid privately to see a psychotherapist and had psychodynamic therapy. This has been the most helpful in terms of trying to repair and understand the damage I experienced as a child. Financially, though, this has been difficult, and I have had to get another job, in addition to my full time job to pay for this.

Depression for me has changed over time, I believe, due to the psychodynamic therapy I have had. For years when I was depressed I needed to sleep a lot and I also put on weight. Now I struggle to sleep (which has its obvious disadvantages) and I tend to lose weight. I didn’t recognise I was depressed for a long while and by the time I went to see my doctor, it was too late to treat successfully, and so took 2 years to recover from. Whereas now it can very quickly become severe, but on a positive note it can ease quickly as well.

Depression is with me all the time, rather like chronic back ache it is always there, but some times are better than others. I have managed to qualify at university in the career I have always wanted, and I love my job, and know that I am pretty good at it. However, there is always the fear that I will get too ill to work. I have had to have the odd day/week off over the last few years, but with the help of my GP (who has been very supportive and allows me to manage my depression my way) I have not had to say it is because of depression. There is a general acceptance at my place of employment about having depression, so long as it doesn't interfere with one's work.

However, I have an excellent manager at work with whom I can be honest. On one occasion I told him that I was going to have to take sick leave as I was very depressed and could not work. He advised me that I could take time off of work, but that if I wanted, he would go through everything I needed to do. He told me that if I felt unable to do something, he would get someone else to do. I went through my work with him, and was able to do everything because he took the pressure off me. He told me to see him at any time I felt unable to do something. Every morning for about a month after that, he would come into my office in the morning to see how I was, and I never took any sick leave.

I have had to build my life around periods of depression, for which I am resentful. I often feel that my life is hanging by a thread – that at any moment, my life, that I have worked so hard to build up, could be taken away from me. It is on this basis that I choose not to engage in a long-term relationship. I am currently seeing some- one, but because of his commitments, I do not see him often. This suits me as it means I am under no obligations or pressure from him.

I feel frustrated that there are no services available to me now. On the surface, I function very well; no one would ever believe that I have depression as I am a good actress. But when it is severe, it would be helpful to be able to access services immediately from a team that knows me and can support me without me having to go through a series of assessments and then being told 'well you can go on the waiting list for this service, but you can only have this service for a particular length of time'. I also feel that long-term psychodynamic therapy should be available, on the NHS, which can get to the root of the issues that cause depression. I now know that I will have depression until I can resolve my childhood issues.

4.2.31 Personal account F

I was first diagnosed with depression in 1999 when I was 44 years old and was feeling suicidal. Because of the way I had been feeling I was relieved to have a diagnosis. Only my close friends knew that I had depression – I didn't want people to know because there is very little understanding within my community.

My mother died when I was 15 years old. My father then attempted suicide and was on a life support machine for 2 weeks. He was brain damaged and I looked after him for 25 years until his death. I was married at 18 and my first child was kidnapped by her father after I left him. My daughter was 3 months old at the time and I never got her back. I married for a second time, to a man who became a violent alcoholic. Because of his drinking he lost a lot of jobs because he was too hung over to turn up and we were often in debt and lived in poverty. We had four children but we could not provide them with much at Christmas and for birthdays. We struggled financially to provide food and the basics.

When I became suicidal I went to see my GP. He was very attentive and took me very seriously and referred me to a psychiatrist and a mental health clinic. Antidepressants and counselling were discussed as possible treatment options and I was referred for counselling but had to wait 18 months, which was useless. I tried various medications, such as Prothiaden, which made me worse. In the end I was put on Prozac which did help to improve my symptoms. When I finally saw a counsellor, I was offered hypnotherapy, which I didn't

want. I wanted counselling. My relationship with my psychiatrist is non-existent. My doctor doesn't have a clue who I am. I'm just another number in a long queue.

I have attended a Christian counselling organisation in the city where I live which has been brilliant. There were well-trained counsellors available who were very supportive. Two of the counsellors maintained contact in between appointments.

Depression devastated my life. I shut out a lot of people because I could not socialise when I was so ill. I didn't want to make relationships because I lost trust in people. My family suffered as I was not really there for them and I couldn't work because my illness was too severe for me to function normally. The house became a tip.

However, things have improved over the years. At the current time I am still on antidepressants but I am ready to come off them. I am now very seldom depressed. After 9 years of being off work because of illness I am now getting back to work on a job placement. If I have any low moods I go back to my counsellor and exercise regularly and eat healthier food to stay well.

4.2.85 Personal account G

I was first diagnosed with depression in 2000 at the age of 42. At the time I was diagnosed, I was unemployed having been made redundant several months previously and also my marriage was in difficulties. I think that these things contributed to triggering my depression but neither was responsible in its own right. On reflection there were signs of problems a couple of years previously.

The diagnosis was not a surprise as it had taken a few months for me to decide to go to see my GP as I tried to cope with it as best as I could. At first my GP was reluctant to do anything but after several visits she relented and prescribed me an antidepressant. Unfortunately, this antidepressant did not work and a few months later I returned to see my GP and asked to see someone. Fortunately my wife at the time had accompanied and backed me up otherwise I don't think the GP would have referred me to a psychologist/psychiatrist.

Initially I had three sessions with a psychologist who said that she could not help and referred me to a psychiatrist. He changed my antidepressant and I then saw him on a monthly basis. This second antidepressant did not work and it was changed again. Eventually I was prescribed a mix of a tricyclic antidepressant and lithium carbonate that proved more effective at controlling the symptoms. However this took 18 months, during which time I was unable to work, my marriage broke up, and because of how I was feeling, I isolated myself from my family. Up until that point I had no experience of mental illness or knew anyone who suffered from it. I was given no information about it from my GP, psychologist or psychiatrist. I think that was the reason I isolated myself from my family more and more as time went on.

During the 8 years I have been ill, I have been on medication and although no longer on lithium I feel that it is only over the last year or so that I have been listened to by my GP and psychiatrist. Since being ill I have changed my GP four times due to moving around the area (one GP retired). Their approach has differed, and has often been inconsistent, and it is only my most recent GP who I feel has listened to me and worked with me dealing with any medical issues around my condition, such as side effects. The one real issue I have about my treatment is that over the 8 years I have only had three sessions with a psychologist and the rest of the time it has been purely medication. I feel this has slowed my recovery and has left me to deal with several issues that I feel could have been dealt with by a psychologist or psychiatrist. Once my condition had stabilised the only contact I had with my GP and psychiatrist was to either get my prescription renewed, or seeing my psychiatrist every 3 months for 10 minutes. Other than that the only other contact I had was with the nurse who took blood samples to check my lithium levels. Also it concerns me that I was never offered

any help or advice on managing my condition. I have obtained such information from what I have discovered on the internet and from fellow service users and the voluntary sector.

As my condition improved I started to research my illness online and also made online contact with others from across the world suffering from mental illness. I have found the internet very useful for getting information about my condition and when I was very ill and needed to talk, I could usually find someone somewhere in the world to talk to 24 hours a day. The other advantage was that when I didn't feel like talking, I didn't have to. Over the years I have formed an online network of fellow sufferers and we keep each other up to date on anything of interest happening in the various countries regarding mental illness and its treatment.

The biggest effect depression has had on my life is when it comes to employment. Since being diagnosed I have only worked for 8 months in paid employment. I've also done voluntary work for 18 months with a variety of organisations involved with disability and mental health. Although I did not have a problem getting work before being diagnosed, since then I have found it difficult. In October 2002 I went to university as part of my 'recovery' graduating with an MSc in 2003. Although this did not help me find work I found it very beneficial to me in that it kept my mind active and this is something I have continued to try and do since then.

Although I feel well at present, it is noticeable to me that my mood is more variable than when I was on lithium, but the strategies I have in place help me cope with this. Also keeping my mind active helps and doing voluntary work gives me a feeling of having 'value' in society. I still have some issues due to the depression, but know that it will take time to resolve these so I try not to let this affect me.

4.3 Personal accounts - carers

4.3.1 Introduction

The methods used for obtaining the carers' accounts was the same as outlined in Section 4.2.1, but for carers of people with depression, the questions included:

- How long have you been a carer of someone with depression?
- How involved are/were you in the treatment plans of the person with depression?
- Were you offered support by the person's practitioners?
- Do you yourself have any mental health problems? If so, were you offered an assessment and treatment by a healthcare professional?
- How would you describe your relationship with the person's practitioner(s)? (GP/community psychiatric nurse/psychiatrist, and so on)
- Did you attend a support group and was this helpful? Did any people close to you help and support you in your role as a carer?
- In what ways has being a carer affected your everyday life (such as schooling, employment and making relationships) and the lives of those close to you?

Two personal accounts from carers of people with depression were received.

4.3.2 Personal account H

Firstly, I must say that caring for someone is one of the most rewarding things I have done. It can be frustrating, exhausting, challenging to one's own physical and mental health, but ultimately helping someone make the most of their lives by helping them in their most vulnerable moments, is rewarding.

This applies to any caring. I was my mother's carer when I was a child and teenager and I made sure she ate properly and took her tablets. But most of all I provided practical and emotional support. But I think it can be damaging for children to care for an adult without support, because childhood is when we should be able to expect to be nurtured ourselves.

I then became a carer to my partner. My partner has had two long periods of depression; at present he has been ill since 2005. They have tried the newer antidepressants on him but one of the old favourites seems to be doing the trick. I attend his reviews and make sure he is looking after himself as regards to diet and exercise. I also emotionally support him by listening, working through problems with him, and trying to encourage him to be positive. His best male friend and I have decided to only respond to positive subjects that he brings up, as a way of trying to create positive thoughts in his repertoire. I have struggled for 2 years to try and get him CBT without success, as I can see he desperately needs to be helped with changing his thought patterns to positive thoughts, which would help his overwhelming depression.

As his carer, the pressure of his overwhelmingly negative thoughts and depressed ways of thinking can be a burden. He doesn't want to think about bills and money, and runs up huge phone bills when he is depressed. I have to constantly nag him to get him to try and keep an eye on his expenditure as it is a risk to his welfare.

As a result of this illness, we can't live together anymore. I see him two or three times a day at either his home or my home, but the pressure of 24-hour depression wasn't doing me any good and I had to move house to be able to care for him again. It actually has the good effect of getting him out of the house at least once a day, to come and see me. I plan trips out, organise things and occasionally exert pressure to get him out of bed and even out of the house, because sometimes he would rather sleep 18 hours a day every day.

His physical health is suffering as a result of extreme weight gain because of the medication and a lowering of his activity levels both because of medication and depression. I battle with his doctor and social worker over this, trying to get them to take this seriously because his father had two strokes at his age and he himself has been warned about fat around his heart. I am trying to get him a review of his medication plus a referral to an occupational therapist for support around physical exercise. It's hard for me seeing him suffer, and sometimes I get angry with his social worker, when they can't see that physical health and other risks are associated with his depression, and that these things should be included in his care plan. It's a constant battle to not get services withdrawn. At one point last year he hadn't seen a social worker or a housing support worker for 3 months, so it's an uphill struggle.

I have neuropathy and sometimes this overwhelms me and I have to lie down for a couple of days to let it 'wear off'. My partner is able to get my shopping and visit me and strangely this seems to take his mind off his own suffering for an hour or two, as he still has physical strength. If it goes on too long, though, he gets cross, and wants me there to support him.

In a way, as a carer, I am more like a mother than a partner, and though I wouldn't say this to him, it has changed the dynamic between us forever. Most carers I have met also say this.

When my partner was depressed previously, I was able to support him and get him back to full time work within a year. Now he has been off work since 2006, and his employers have given him until December 2009 to get through this depression, but I know it is a real risk for him and not working in the long run would not help his self-esteem.

I have built my career around being self-employed, and working from home in the mental health and housing fields, mostly regarding carer, resident or service user issues at strategic level. This means I have the time to care, but I am able to keep myself busy and to have time for myself through work. Work is very, very important to most carers: I have heard other carers say that they go to work to get a rest from the overwhelming nature of caring.

The role of being a carer for someone with severe depression has added to my own symptoms of dysthymia over the years because of the sheer pressure of coping with someone who turned down treatment, stopped their antidepressants at one point and crashed into a psychotic depression. This was a huge burden and local services left me to cope with this on my own 24 hours a day, and it nearly broke me.

Carers who become ill with depression or anxiety, or who have a previous history of depression, should be offered support. As I have said, caring is rewarding but it can also be tiring and frustrating.

4.3.39 Personal account I

My Mum has been depressed on and off since I was a 7-year-old boy (I am now 15) and I have been caring for her since then. She's not depressed all of the time, and it's fun when she's well, and normal, like – we do normal things then and she's the normal bossy Mum.

When I was small it was just making her a cuppa now and again, or telling her about school with funny bits to try and make her laugh. Or telling my Nan and Grandad about how she was so they could come and help, but now it's more. I sit down and talk with her, make sure I get in straight away from school because I worry about her when I am out. I get her tablets, make appointments, sort out food shopping, nag her to get dressed when she's depressed, and answer the phone. I am more of a grown-up than when she's well.

Mostly she's well but now and again she gets depression. I know the signs. Then she goes quiet and stops going out and seeing her friends and I try and cheer her up and make things better for her. I wish she was like other Mums sometimes, and, well, all the time. But I wouldn't be without her or want to leave her on her own – she's my Mum! I try and be positive and jokey, behave myself and be there for her, and make sure she sees her therapist even when she doesn't want to go out and sometimes get her friends around for a surprise to make time pass for her. I hope she gets better soon. I go to my room when I feel cross and sometimes talk to my friends. I go out and do usual things too so that she doesn't worry about me. I do well in school.

My Mum takes tablets and sees her therapist but I think seeing people really helps her. When her friends come round and take her mind off it for a while, she laughs. Don't forget your friends when they are depressed, I say. And chocolate sometimes helps too!

For a while I had no support but now I go to the Young Carers' Centre in our town, and I meet other people like me caring for their parents. I play pool and we have days out – we went to Alton Towers which was fun. It's good meeting other young people like myself who are carers too, but we don't talk about it all the time. We want to get away from it just for a few hours, fool about, be normal. Sometimes we watch films, have pizza, and there's a support worker if you do want to chat. I had a carer's assessment there too. People sometimes think or say my life is sad, but I know it's not my Mum's fault, she can't help being depressed. I love her and where else would I want to be? She helps me too.

4.4 Qualitative analysis

4.4.1 Introduction

The following section consists of a qualitative analysis of personal accounts of people with depression using Healthtalkonline (www.healthtalkonline.org). Healthtalkonline provides interviews with people with both physical illnesses and mental health problems. The review team undertook their own content analysis of the interviews to explore themes that could be used to inform recommendations for the provision of care for people with depression.

The same transcripts were also reviewed by Ridge and Ziebland (2006), which is included in the review of the qualitative literature below. The review team decided to undertake their own analysis to cover a wider range of themes than those focused upon by Ridge and Ziebland.

4.4.24 Methods

Using the interviews available from Healthtalkonline, the review team analysed the experience of 38 patients from across the UK. The methods adopted by Healthtalkonline to collect interviews were two fold. First, the participants were typically asked to describe everything that had happened to them since they first suspected a problem. The researchers tried not to interrupt the interviewees, to obtain a relatively unstructured, narrative dataset. Second, a semi-structured interview was conducted in which the researcher asked about particular issues that were not mentioned in the unstructured narrative but were of interest to the research team.

From the interviews, the review team for this guideline identified emergent themes relevant to the experience of people with depression that could inform the guideline. Each transcript was read and re-read, and sections of the text were collected under different headings using a qualitative software program (NVivo). Two reviewers independently coded the data and all themes were discussed to generate a list of the main themes. The anticipated headings included: 'the experience of depression, 'psychosocial interventions', 'pharmacological interventions' and 'healthcare professionals'. The headings that emerged from the data were: 'coping mechanisms', 'accessing help and getting a diagnosis of depression', 'stigma and telling people about depression' and 'electroconvulsive therapy'.

There are some limitations to the qualitative analysis of people's experience of depression and its management undertaken for this guideline. As the review team relied on transcripts collected by other researchers with their own aims and purposes, information on issues that are particularly pertinent for people with depression that could be used to inform recommendations may not have been collected. Moreover, the review team did not have access to the full interview transcripts and therefore had a selective snapshot of people's experience. However, using Healthtalkonline did highlight issues regarding depression that can be reflected upon for the purpose of this guideline.

4.4.30 Experience of depression

In recounting their experience of depression, some people described life events which they felt had caused the disorder. Some of these events were childhood experiences including both problems in the family and at school. Some people commented that stressful situations at work contributed to the onset of their depression. Many people described the death of a family member or friend as a trigger of their depression. One service user summed up various life events that she believed were associated with her current state of depression:

> 'All these experiences from earlier on in life, my Mum dying, being bullied … being neglected and isolated and being treated different academically. I think they all combined with my lack of social skills, which I'd not had a chance to develop until that point when I got to university … within a few months … I was just feeling very low and very lonely, needy … I think, probably about 4 or 5 months after starting my first year, I did become very depressed.'

Some people used metaphor and allusion to illuminate their experience of having depression. For example, one person described having a 'racing' mind that was 'zooming into miserable places'. Others used analogies such as depression being like a 'brick wall' or 'being inside a balloon' to describe how depression can act as a barrier from experiencing the world:

> ‘I couldn’t feel anything. I couldn’t feel anything for [husband’s name]. I couldn’t feel anything for the children. It [depression] was like being inside a very, very thick balloon and no matter how hard I pushed out, the momentum of the skin of the balloon would just push me back in.’

Other people listed the symptoms they were experiencing: lack of pleasurable experiences, body aches, tearfulness and sleep problems; they also described feelings of loneliness, isolation and feeling withdrawn.

A prevalent theme in the interviews was the presence of negative thoughts. These thoughts were described by people with depression as irrational and often caused them to jump to conclusions. One person explains how she experienced negative thoughts:

> ‘I call, what I’ve got in my head my chatter box. Basically it is my mind, seeing things a particular way. And with depression you see it really negatively. You see everything negatively, you’ll always pull out the negative over the positive if you ever see a positive, you’ll … if for one positive you’ll give ten negatives.’

People also described feelings of suicidal ideation and some disclosed their experiences of attempting suicide. Some of the suicidal thoughts relating to suicide were: the ‘world would be a better place without me’, ‘life wasn’t worth going on’, and ‘life was completely out of my control’. One person described a suicide attempt:

> ‘I can remember being almost unconscious, and with a doctor and nurses around the bed. And the doctor said to one of the nurses, ‘Go and get so and so … we’ve got about 10 minutes or he’ll be gone’. And I could hear him, and I just thought, ‘I wish you’d leave me alone. I’m warm and comfortable. I don’t want this.’

However many people also identified positive aspects of having experienced depression, for example, having become more confident, positive, understanding of others, able to support others and able to do ‘something positive and … creative’. They also said that they had become more aware of themselves and their feelings and more able to cope with stressful events.

Another common theme was that people felt that they appreciated life in a different way after having been depressed. For example, one person said:

> ‘I can listen to music and appreciate it in a different way … it can move me now. Something on the TV can move me now, and I have, I feel things and things affect me.’

Many people also felt that experiencing depression had made them re-evaluate their lifestyle and that this had led them to make some important positive life changes. One person described having had a breakdown as a ‘breakthrough’. Another person described the positive effects of having had depression:

> ‘I think it’s [depression has] sort of made me question what I thought was good about my life because I was in a very busy and hard-working career, and whilst the depression wasn’t the main, or the only reason, that I left, there was a re-organisation at my work, I do think, oh, thank God I left there when I was 36 rather than 56. You know, I understand that I need sort of time for me now, and that I’m a person in my own right, and I’m important and I have, you know, the right to have some quality time for me.’

4.4.4.4 Accessing help and getting a diagnosis of depression

Some people detailed how a particular event or problem prompted them to access help, such as sleep deprivation and lack of concentration:

'I was putting my eldest daughter to bed and trying to read her a child's story, and I actually found … I no longer had the concentration to read … I couldn't follow the sentences to actually read it out loud. And that was a point where it was clear that … I had to seek help. And so I made an appointment with the doctor the next day.'

Once people with depression accessed help, they described their experience of receiving a diagnosis of depression. Some described how there is not enough recognition of depression and how often when they presented with sleep problems or loss of interest in sexual activities to their GP, these symptoms were not initially recognised as symptoms of depression:

'I went to the doctor and I said … 'I sleep but I always feel tired … I've tried … everything.' And he just said, 'Try getting more sleep.' [laughing] I was like, yes, I could have thought of that, I've tried that, it didn't work … my feeling is that really he should have asked a few questions and could possibly have diagnosed that I was depressed.'

4.4.5 Stigma and telling people about depression

Some people described the stigma of having a diagnosis of depression. The majority felt that stigma still existed while a minority thought it was less prevalent than it used to be. There was also stigma around receiving treatment for depression for both psychological and pharmacological interventions:

'It took a hell of a lot for me to go to therapy. You know A: nutters go to therapy, B: therapy makes you a nutter. These were the kind of things that I grew up with. And it doesn't help. You know, so hostile kind of lower middle class sort of feeling about that sort of thing.'

Conversely one person said it was quite 'fashionable' to be taking medication:

'Prozac is quite a fashionable antidepressant. And it was OK to say you were on Prozac, it's like a happy pill isn't it. I'm OK I'm taking Prozac and then of course I knew quite a few people who were taking it as well, so it was like ok like join the club.'

Due to the stigma surrounding depression, some people found it difficult to talk to other people about their condition:

'I can't talk to my family about it. They don't know about the therapy. I think it's the stigma thing … my perception is that I would be seen as weak and not coping, so it's easier for me not to admit to that weakness.'

However, some people encouraged others to speak openly about their condition:

'You should tell someone now, it doesn't have to be the doctor or a therapist, it can be a friend you know. The older I've got, the more I've found that it's acceptable to say to people, "I'm depressed at the moment".'

Some described their experiences of telling friends and neighbours and stating that it helped them; one person made a joke to ease the situation:

'I was just really outright, and I just said, "Ok, I was in a psychiatric hospital for a month and then outpatients for a further month and now I'm at work part-time to try and get back into the swing of things slowly." And he just looked at me … I said, "It's ok though," I said, "I'm not loopy" and he just started laughing, because I'd just turned it into a joke.'

4.4.6 Psychosocial interventions

People with depression discussed their positive attitudes towards psychological treatments:

> 'Sometimes you do need to talk to somebody who you don't know, who understands, instead of chatting to the brick wall. And instead of it going round in your head and trying to sort it out. Or you need somebody to talk to you and push the right buttons to help sort yourself out.'

People with depression expressed the need for psychosocial interventions when the cause of depression was deemed to be psychological rather than a 'chemical imbalance'. In addition they explained how they thought psychosocial interventions, rather than medication, were needed to resolve the maladaptive behaviour and distorted thoughts that contributed to their depression:

> 'These tablets helped me … but after a while, I realised it sorted out my brain chemistry, but you have learnt all these negative ways of looking at things, and doing things … and that is why I believe I need long term therapy as well. I felt better [with medication], but I still didn't have ways of dealing with things.'

The benefit of psychosocial interventions to tackle negative thoughts was a prevalent theme. People described how they learnt to change their thoughts to be more constructive and positive:

> 'There are things that keep me in a place of being depressed, and … that's what the therapy really helps … me understand how I perpetuate the depression … I think for me it's about blaming myself … thinking that I'm a bad person, and I can expend huge amounts of energy on the mental processes that go into making me responsible for everything that goes wrong in the world.'

In the following sections, experiences of different psychosocial interventions are described by people with depression. The psychosocial interventions that were briefly touched upon were counselling, cognitive therapy, self-help material, relaxation therapy and support groups.

Counselling

Overall people who discussed having counselling were positive about their experiences:

> 'The main sort of release point was the counselling, which to me was crucial. If I hadn't have had the counselling, I'd probably still be severely ill and wouldn't be, you know, happily now saying that at last I'm enjoying life to a greater extent.'

Some of the outcomes that people achieved from counselling were: an increase in self-esteem, being able to return to work, dealing with bereavement issues, learning more about oneself and helping to deal with thoughts and feelings. Counselling was a positive experience for many because it provided a safe environment in which to talk about their concerns:

> 'It was a big relief to have someone who I could tell anything I wanted, anything that was bothering me, and not worry about what they might think about it or how it might affect our relationship. And you know, it also helped to feel that I was doing something about my problems as well.'

Cognitive therapy

People who had cognitive therapy were positive about it, describing it as enabling because it was practical, focused on the real world and allowed them to begin to help themselves:

> 'I could change my thinking and I could thereby change my feeling … A particular example was he [therapist] said, when you go lie down to go to sleep, he said, "You tend to look back on your day and think of all the failures" … "why don't you just think of everything that's been successful?" So … I started doing that … So just things like

that, a few things like that with cognitive therapy. You know I think they helped quite a bit.'

Self-help

Two people described using self-help books to cope with their depression. One read David Burns' Feeling Good, which is based on cognitive and behavioural principles:

'I sat and read this book, and you know it's quite a hefty one. But it's a really good one …. It's very difficult to sort of … stop yourself, and realise that just because you have an opinion or you express yourself a certain way, it's not right or wrong, to you know, to act that way … it's really difficult, 'cos it's everything in the book ties up with other things and you know cognitive therapy for me, is my chatter box and arguing with it.'

Another read Dorothy Rowe's Depression: The Way out of Your Prison:

'Some of it is relevant, some of it is not at all relevant … It's really good because it's all about … looking after you and some of the things just make me laugh. You know because it's so like … "That's me. I'm in there. That's what I do".'

Relaxation therapy

Two service users described their experience of relaxation therapy:

'Relaxation therapy … when you're depressed is mighty hard to get started. Once you've started and got the grasp of it, then it's quite good, but to actually get relaxed when you're really depressed is damn nigh impossible you know.'

Support groups

People who had attended support groups were positive about their experiences. They described these groups as therapeutic because they were able to meet people with similar problems and share their experiences in an environment where there was no stigma. In addition, people with depression felt relieved to know they were not alone:

'It was a great source of comfort … And to find that in fact you weren't the only person to feel like that was actually a great relief. It was also a great relief to find . . . people who were non-judgemental.'

'A self-help group isn't group therapy but it is very therapeutic … people meeting with a shared interest … There are people there who, they won't say, 'Pull yourself together, pull your socks up, what have you got to be depressed about?' There is none of that. The mutual support is just unbelievable.'

One described a suicide support group that provided some source of comfort but also had harmful effects:

'It's a discussion group of people talking … of essentially extremely depressed people talking about suicide. And talking about suicidal feelings and suicidal methods and yeah, from time to time people die on it. But in a weird perverse way it's a source of strength and a source of comfort.'

4.4.7 Pharmacological interventions

People with depression had mixed views regarding pharmacological interventions. Some people were concerned about taking tablets; they did not think pills solved the problem or they had a cynical view of drug companies. Others who tried medication who did not have

positive experiences said they felt that it 'robbed' them of feelings. One person described why a pharmacological intervention was not the right treatment for him:

> 'I've been prescribed antidepressants in the past but I've always felt reluctant and apprehensive about taking it, largely because a) I feel that the effects are probably short-term, they're not going to actually resolve the depression, b) because they do have side-effects and, c) I didn't feel comfortable, myself, with taking some tablets.'

However, the majority had positive experiences regarding medication. For those who benefited from a pharmacological intervention, they described taking medication as a turning point in their lives. People said that they felt more in control and had greater awareness of the world around them (this was in contrast to other people's experience of medication):

> 'It was exactly 7 weeks to the day that I took … the first tablet … I knew that morning when I woke up that I feel differently, things are different. And that was the turning point. It was this lifting again, this lifting of overall and just … contentedness.'

> It [medication] gave me a feeling that I've got some control now of this thing [depression]. And I was having some experiences like increased sensitivity to things like noise and colours and feelings.'

One person advised that if someone was not benefiting from their current medication, that they should persevere until they found a drug that works for them:

> 'It isn't a one size fits all … I would say to folk if you feel like you're not getting any better … on the particular medication … go back to your doctor and ask your doctor to change, to consider changing your medication.'

Many people with depression reported side effects from taking medication, notably dry mouth, hair loss, increased sweating, weight gain and problems ejaculating. A minority also reported experiencing suicidal thoughts as a consequence of their medication:

> 'For many years I hadn't had any suicide thoughts at all, and I had certainly never thought of cutting myself, but while I was on Seroxat, I did start to get sudden images in my head of you know, cutting long gashes in myself.'

Despite this, some people with depression said that the benefits of medication outweighed the potential side effects:

> 'You're given a sheet which tells you what to expect, and I looked it up on the internet as well. I'm very against taking medicine for a long time, but after my experience with the depression I decided I would be prepared to take it…for the rest of my life if I don't get it again, the depression again, if it stops that.'

When some people stopped their medication, they described experiencing discontinuation symptoms, the most prevalent symptom of which was nausea:

> 'Being stupidly pig-headed, just stopped it (Efexor) … I was just completely off my head with depression … the symptoms were so acute it was very frightening. You feel sick, nausea, the nausea was awful. And just panic, really.'

4.4.8 Electroconvulsive therapy

Four service users recounted their experience of ECT; the majority had negative experiences because of the frightening nature of the intervention and loss of memory post-treatment:

> 'They'd get you to lie down on the bed, and give you an anaesthetic in your hand, which would basically make you go unconscious. But just that 2 minutes when you might have gone into the room and been waiting, I was just so frightened. And then

> they give you ECT … that is quite a confusing experience. I did find that it affected my memory a fair bit.'

> 'I have massive blanks, short-term and long-term … I get angry with the professionals that this wasn't explained that this could happen … I've tried to talk about it with the doctors at the hospital and they say, "Give me an example" and I give them an example and they say, "Oh that's normal, that's just normal, that's not the ECT … that's normal".'

Only one person reported a positive experience regarding ECT:

> 'It all sounds very scary, but you really don't … you don't see anything because you are anaesthetised, so you are asleep. And you wake up, and I … you have a slight headache, but apart from that, I had no side-effects … my mood improved instantly, and I was talking and laughing.'

4.4.9 Healthcare professionals

This section covers people's experience of healthcare professionals, including GPs, nurses and psychiatrists.

GPs

As described in Section 4.4.4, people were critical of their GPs because they felt that their depression went undetected. However some people had positive experiences of getting a diagnosis of depression and of how their depression was initially managed:

> 'I was very low physically and clearly very low mentally, and the GP … and I'll be forever thankful for him, actually said, "I don't think I am helping with the right kind of medication for the right reasons, and if you agree I'd like to refer you on to somebody". And it was like an immense relief … somebody's actually going to treat me as somebody who has a problem here.'

People who had positive experiences of their GPs described them as being sympathetic, warm, tender, kind, helpful and supportive. These people felt that they were listened to and responded to:

> 'She's [the GP is] good because she is human. She listens and she responds to me as a human being, not as a professional. She gives me time, as much time as I want sometimes. She cares and she's shown me she cares because she has rung me up before at home and said, 'How are you? Will you come and see me tomorrow?' because she knows I'm not going to ring and make an appointment because I … I mean I'm in isolating mode and things are going wrong.'

Those with negative experiences described how their GP was lacking in the above characteristics:

> 'You just didn't get listened to, you didn't get, you know, it was as though what they [GPs] were saying was, "Well, it's just in your head, you know you don't really understand, I know better." And I know that they're really busy and I know that they don't have a lot of time, but I really felt that I got no help at all most of the time.'

Nurses

People said that they did not feel that nurses understood the sensitive nature of their depression, that nurses in the NHS were too busy to talk to their patients and that their attitudes may be because of inadequate training:

> ‘There’s an awful lot there who … you felt as though it was people saying to you, “Oh, for goodness sake pull yourself out of it”, and, “Get yourself together”, which you don’t want, it’s the last thing at the end of the day. I just don’t think that there is enough, in regards to, against private and NHS, there is just not enough funding to be able to … I don’t know, train the nurses in a certain way.’

Psychiatrists

People had mixed experience of psychiatrists. Some did not like how psychiatrists tried to illicit information about their childhood experiences, describing the method as a ‘text book’ approach that instantly created a barrier. Others did not like to discuss feelings in general:

> ‘I felt my psychiatrist was a very … oh … wet individual. Again, I think because I’d been quite a numerate, factual, organised person, to have someone to talking about feelings and what about this and what about that? And it was … nothing could ever be pin-pointed or … I just found it annoying.’

People also had mixed opinions about how their psychiatrist dealt with their medication. The majority had positive experiences: one person described how their psychiatrist was able to change their medication to one with fewer side effects; another described how the psychiatrist prescribed a proper therapeutic dose of anti-depressants. However, one person felt that she was not listened to when she explained to her psychiatrist that her current medication was not working:

> ‘He’d [psychiatrist] say something like, “Oh well, continue with the paroxetine.” And if I said, “Look, this isn’t helping me. I’ve been on this for eight months, it’s not making me better.” “It takes time, you have to have patience.” You know, “You are better really” I was told by one doctor. “You’re not depressed, you’re just a very sad lady.”

4.4.10.4 Services

The experiences of mental health services were described by people with depression. Issues regarding referral, waiting lists and getting into NHS services were raised. Some people said that that they waited too long to be referred to a psychiatrist or receive psychotherapy. One person said that while she was on a waiting list she was unable to cope with her depression:

> ‘I was referred to the psychiatric hospital for assessment. Although I think it probably took about two months I believe between the initial sort of GP’s referring letter and getting an appointment. Which again in retrospect was, was way, way too long, way too long. I was really, really ill and barely coping.’

Another person described how she felt that she had to be violent in her GP’s surgery in order to be referred to NHS services:

> ‘It’s very difficult to get a hospital bed for quite severe mental illness. You’ve got to be suicidal … I was feeling suicidal. I was also quite violent at times. I mean in my own doctor’s surgery, I swept all the things off his desk you know … there was a part of me, kind of watching what I was doing … saying, “Right, well make it really dramatic.” I wasn’t pretending exactly, but I knew I had to make a song and dance to get heard.’

Once in mental health services, people described a mixture of positive and negative experiences. One person said that a psychiatric intensive care unit was ‘a place of safety’. Others described a mental health service as a place where they had no responsibilities, where they could ‘hand yourself over’ to the care of the service. Accompanying this, however, was the feeling of being institutionalised:

> ‘In eight weeks, I very quickly became institutionalised myself. I was scared to come out because I was in this enclosed world where I knew what was going to happen.

> There were routines, mealtimes, getting up times, medication times, OT (occupational therapy) times. There were routines and I had no responsibilities … I was in a place where I didn't have to think about anything, and nobody could touch me.'

People also had negative experiences of mental health services provided by the NHS, including not feeling cared for. Those who had had private treatment had more favourable accounts, and compared and contrasted the two experiences:

> 'The private hospital was, there was a lot of love, a lot of care in there, sincere care. And I won't knock the NHS because they are obviously very limited to money in a way, but there was no care … In the private hospital you felt like you were being treated as a human being … You felt that yes, you could get well here because they cared.'

4.4.11 Families and carers

People with depression described the impact that their condition had on families and carers. Some stated that it was harder for the family and carers than it was for the person who had depression. Others described the impact that it had on the partner, often resulting in a change in roles. For example, people described how their partners had to take a more active role in daily chores:

> 'I found it difficult to relate on the day-to-day things, which is where she (his wife) was so good. She took over those things.'

Some felt that their depression had an impact on their children:

> 'My sons were very good, but they missed a lot because of how I was. And they would have to make allowances, which isn't really what you should have to do when you're growing up.'

Some people said that without their family and carers they would not have been able to cope with their depression:

> 'My partner has played a key role in my recovery – he was very supportive during my depression periods – I do not know how I would have coped without him … Many times he has forced me to do things and helped me out of the house in times when I did not feel like doing anything. I believe having a loving and caring partner has helped me get over the most horrible periods of my depression.'

4.4.12 Coping strategies

People with depression described coping strategies that they used to overcome their condition. These strategies were those other than pharmacological and psychological interventions employed by people to manage their depression.

Distraction was a common coping strategy. One of the ways in which people distracted themselves from their mental health problem was by having or acquiring a hobby, which ranged from physical activities such as swimming and going to the gym, to those of a more creative nature such as poetry:

> 'Having hobbies, and that … that gets depressed people through because the thing that you can't think of, you know, two things at once.'

> 'I wanted to do something physical … So I started to garden, I've never been in the garden before. And it was crap at first, but gradually it was alright, you know you start to think, "Yeah, this is kind of distracting me a bit."'

For other people, voluntary work was a coping strategy because the process of helping others allowed them to help themselves. In addition, people described how voluntary work helped them to increase their confidence and build up their self-esteem:

> 'At the beginning I used to get anxiety attacks and some days I could just phone up and say, "Look I'm not feeling well." If you are doing it voluntarily … I felt I wasn't letting them down … the same pressure is not there. So … voluntary work I would definitely advocate because it gives you a sense of … it helps build your confidence, self-esteem.'

Another coping strategy was completing small, manageable tasks:

> 'When I'm depressed … I wasn't able to do anything about it, really. I just felt overwhelmed by it … And with my depression, when I was feeling very low, I would, I did decide to just concentrate on small things; going for a walk, baking some bread, you know pottering around in the garden. Just trying to get through day to day, I think, was how I came out of the suicide attempt.

4.5 Review of the qualitative literature

4.5.1 Introduction

A systematic search for published reviews of relevant qualitative studies of people with depression was undertaken. The aim of the review was to explore the experience of care for people with depression and their families and carers in terms of the broad topics of receiving the diagnosis, accessing services and having treatment.

4.5.2 Databases searched and inclusion/exclusion criteria

Reviews were sought of qualitative studies that used relevant first-hand experiences of people with depression and families/carers. The GDG did not specify a particular outcome. Instead, the review was concerned with any narrative data that highlighted the experience of care. For more information about the databases searched see Table 7. Details of the search strings used are in Appendix H.

Table 7: Databases searched and inclusion/exclusion criteria for clinical evidence

Electronic databases	CINAHL, EMBASE, MEDLINE, PsycINFO, HMIC, PsycEXTRA, PsycBOOKS
Date searched	Database inception to February 2009
Study design	Systematic reviews of qualitative studies, surveys, observational studies
Population	People with depression and families/carers
Outcomes	None specified

4.5.3 Studies considered

The search found one systematic review that explored the experience of care for people with depression that met the inclusion/exclusion criteria (Khan et al., 2007). The review team then looked at primary qualitative studies identified by the search and a further two primary studies (Ridge & Ziebland, 2006; Saver et al., 2007) were included in the review that were not already reviewed by Khan and colleagues (2007). A further seven studies were considered for the review but they did not meet the inclusion criteria (Cooper-Patrick et al., 1997; Rogers et al., 2001; Chew-Graham et al., 2002; Van Schaik et al., 2004; MaGPIe, 2005b; Elgie, 2006; Johnston et al., 2007); the most common reasons for exclusion were the studies did not report qualitative data or the population did not meet criteria for depression.

4.5.41 Themes emerging from the studies

2 Experiencing depression

3 Khan and colleagues (2007), in their meta-synthesis of qualitative research in guided self-
4 help in primary care mental health services, found that family conflict, problems at work,
5 chronic physical health problems, childhood events, financial hardship and racism were the
6 most frequent reasons given for causes for depression. People taking part in the studies
7 spoke about their depression in terms of the effect on functioning and ability to cope rather
8 than feelings or symptoms. The most common means of expressing their feelings was
9 through metaphor: being 'on edge', 'boxed in', 'a volcano bursting', 'broken in half', 'prisoner
10 in my own home', and so on.

11 Accessing help and stigma

12 Khan and colleagues (2007) found that accessing help from primary care could be difficult,
13 with very little time spent having one-to-one contact with a primary care professional.
14 Because of feelings of shame and 'lack of legitimacy', people may not have presented their
15 problems in an open manner. There was a possibility that seeking help would 'threaten an
16 already weakened sense of self' if treatments were discussed that might be unacceptable to
17 the person, such as medication.

18 Saver and colleagues (2007) described four barriers to accessing help by people with
19 depression. These were characterised as: (1) a lack of motivation because of their
20 depression; (2) stigma associated with depression and/or denial of their diagnosis; (3)
21 healthcare professionals seeming unresponsive; and (4) a mismatch between how
22 information is offered and how people with depression prefer to seek information, for
23 example:

24 > 'I would never sit down and read something about medicine. It has never interested
25 > me. I learned more from watching that commercial on television.'

26 Getting a diagnosis of depression

27 For people with depression, Saver and colleagues (2007) found that the majority of people
28 received their initial diagnosis from a mental healthcare professional and a minority reported
29 receiving their diagnosis from a GP. In addition, people said that their GP missed
30 opportunities to diagnose their depression. Some people described their own inability or
31 unwillingness to raise the issue of depression with their GP, while others stated that their GP
32 focused solely on their somatic complaints, seemed uninterested in mental health issues or
33 were purely dismissive of depression when it was suggested.

34 Experience of treatment

35 Khan and colleagues (2007) found that taking medication could lead to ambivalent feelings:
36 on the one hand, people felt relief because medication helped them cope with difficulties in
37 their day-to-day life; on the other hand, they felt a lack of control. There was also a moral
38 component regarding personal responsibility and the fear of not being able to function in daily
39 life. When the GP or others (family or friends) offered advice to relieve this ambiguity, people
40 were more willing to accept medication as a possible treatment, but only on the
41 understanding that it would be for short-term use. People were cautious about telling other
42 people that they were taking medication because of perceived stigma. There was a feeling
43 among the people in the studies that they were in some way 'deficient' because they needed
44 to take antidepressants. Feelings of guilt, of letting themselves and others down, and
45 concerns about long-term changes to their personality were also expressed.

Saver and colleagues (2007) found that less than half of the people with depression reported receiving information about psychological interventions. One participant commented that the only 'option' was a pharmacological treatment:

> They just handed me a drug and said go on it right now … I felt rushed along, given a prescription, told this will fix it.

None remembered receiving information about the different treatment options such as CBT, problem-solving therapy or IPT. Only a minority reported that they had some choice in their treatment options.

Ridge and Ziebland (2006) in their analysis of interview transcripts collected by Healthtalkonline found that people with deep-seated and complex problems needed longer-term psychological therapy.

Self-help and other coping strategies

Khan and colleagues (2007) synthesised qualitative studies of patient experiences of depression management in primary care to develop a framework for a guided self-help intervention with the aim of providing a potential solution to the problem of the gap between demand for CBT and supply of trained therapists. A number of themes were highlighted, including feelings of control and helplessness in engaging with treatment, which might influence the success of a self-help intervention for people with depression in primary care. People said that they used coping strategies such as distraction or thinking of places that were associated with feeling safe and in control. They saw accessing help as an indication that their personal coping strategies had failed.

Recovery

Ridge and Ziebland (2006) analysed the interview transcripts (collected by Healthtalkonline) of 38 men and women who, in the main, had had severe depression, to explore the approaches and meanings attributed to overcoming depression. The focus was on the specific components involved in recovery: authenticity, responsibility and 'rewriting depression into the self'. Recovery involved the need to understand the 'authentic self'. The main findings of the study were that people needed to understand a language and framework of longer-term recovery to tell their own story of improvement; that getting better meant different things to different people; and that people needed to assume responsibility for their own recovery. The majority of the interviewees had used and valued talking therapies as a means of gaining insight into their thoughts and feelings.

4.6 From evidence to recommendations

This section is a combined summary of themes from the personal accounts, the qualitative analysis and the literature review. It should be noted that most of the personal accounts received were from people who either have or have had severe and/or chronic depression. Therefore, it is acknowledged that the themes that run through the personal accounts may not be applicable to people who have other forms of depression. Despite these limitations, a number of themes were identified that were present in all three sources of evidence.

4.6.1 Understanding depression

Both the personal accounts and the literature reveal that lack of information from professionals is a barrier to coming to a full understanding of depression, the range of treatments available and the role of the mental health team. There was also a concern that when a person is severely depressed they may find it difficult to concentrate on what is being said. Therefore written information is crucial, although it should be recognised that people with mental health problems may respond to information provided in other forms, such as via

video or DVD. One person (B) said that it would be helpful if professionals could be clear about the purpose of any appointments offered. Lack of clarity about how care is organised may increase the person's distress. One person (G), who had been given no information, had empowered himself through the internet and had built up a wide network of fellow sufferers. Lack of accessible information is a particular issue for people from black and Asian minority ethnic groups, as evidenced by personal account C.

4.6.27 Accessing help and getting a diagnosis of depression

Accessing help was also a prevalent theme in the personal accounts, the qualitative analysis and the literature, whether it was during the initial stages of being diagnosed or after years of having treatment. Two people in the personal accounts (B and E) found it difficult to access support when needed, despite having had depression for some years. It was felt that an emergency number to call would be a lifeline for people who live alone and have no carer support. Such means of support would be particularly helpful for people with long-term, severe depression.

The literature also revealed that accessing help may be a problem for some people first experiencing symptoms because of stigma associated with having a mental health problem (see Section 4.6.3), which may leave them unmotivated to raise the issue of depression with their GP.

4.6.39 Stigma

Stigma was frequently discussed in the personal accounts, the qualitative analysis and in the literature. This was experienced both externally and internally. External stigma was felt from employers and colleagues; but many also felt internal stigma and kept their depression concealed from friends, family and work associates. Feelings of shame were expressed and also an anxiety that asking for help would lead to being offered interventions that they did not want, such as medication (the person in account D said that the idea of taking tablets accentuated the feeling of being mentally unwell).

4.6.47 Recognising depression

Recognition of depression and the severity of symptoms was also a prominent theme in the three forms of evidence. In the literature and qualitative analysis, people spoke about how depression is often not recognised and that physical problems may mask the depressive symptoms or may not be seen as part of the depressive symptomatology. In the personal accounts, two people (B and G) commented that they felt that the severity of their depression was not properly recognised within primary care. One person (B) felt that her diagnosis should have been made by a qualified and experienced professional.

4.6.55 Relationships with healthcare professionals

The relationship with the GP was a prevalent theme in the personal accounts, the qualitative analysis and the literature. In the personal accounts, most found their GPs helpful and understanding. The main area of criticism concerned the quality of contact with the GP (see Khan et al., 2007) – a short appointment when a person is distressed is not long enough and people with depression are unlikely to ask for a longer appointment. In the qualitative analysis and the literature, the relationship with the GP was seen negatively if the GP failed to recognise depressive symptoms or focused solely on the person's somatic symptoms. People who had positive experiences highlighted the sympathetic, supportive and helpful qualities of the GP.

The relationship with nurses was not as positive in both the personal accounts (see B) and the qualitative analysis, with lack of understanding about depression being cited as a common complaint.

In the qualitative analysis there were mixed views about psychiatrists, particularly in the way that they prescribed medication. Some people felt that their psychiatrist was able to work with them to find the right medication and the correct dose; another said her psychiatrist did not listen when she said her medication was not working. In the personal accounts, some people had neutral views about their psychiatrist while three people (C, F and G) expressed negative views, such as the psychiatrist being unsupportive and cursory in their attention.

Most of the personal accounts spoke of the importance of a relationship with professionals that was non-judgemental and supportive. But as one person (B) pointed out, sometimes being well-meaning and supportive is not enough. She felt that while her primary care practitioners and counsellors were pleasant and accommodating, her self-report was not listened to closely enough and the severity of her depression was underestimated. A number of people commented that the relationship between patient and therapist is of prime importance, and that ideally there should be some choice in terms of the gender of the therapist and their therapeutic approach. Two people (A and B) commented that it is often seen as the patient's 'fault' if they do not benefit from psychological treatment, when the counsellor or therapist should take some responsibility for a lack of therapeutic effect.

4.6.6 Experience of services

Both the personal accounts and the qualitative analysis described experiences of mental health services. Many people said that they waited too long to be referred to a psychiatrist or receive psychological treatment. Once in mental health services, views were mixed. In both sources of evidence, those who had private treatment had, on the whole, more positive experiences.

4.6.7 Experience of depression and its possible causes

In both the personal accounts and the qualitative analysis, people with depression described some of the negative thoughts that they had experienced and some described suicidal thoughts and behaviour; they also used metaphor and allusion to explain their symptoms. In the qualitative analysis some people said that they were able to experience life differently since being depressed which, for some, was a positive outcome.

It emerged from the qualitative analysis that some people ascribed the onset of their depression to certain life events, including childhood experiences. The majority of the personal accounts also reported childhood events such as trauma, abuse or conflict of one form or another and many of them linked this directly with the onset of their depression. For many people, complex problems in childhood were compounded by multiple difficulties in adulthood. For the person in account D, being a carer of someone with schizophrenia meant that he had to hide his symptoms of depression to fulfil his role as a carer. Khan and colleagues (2007) found that family conflict and childhood events were among the most frequent reasons given for causes for depression. Howe (1995) explains that:

> 'Internal psychological states and our ability to cope with the external demands of life have roots which reach right back into childhood. The robustness of our early internal representations of self and others lays down the pattern of our future psychological strengths and weaknesses. When children feel that no matter what they think, say or do, they are not able to control what happens to them, physically or emotionally, a feeling of fatalism and helplessness sets in. Attachment relationships in which sexual or physical abuse took place often leave the individual with feelings of passivity and worthlessness. Early attachment relationships that were lost or broken leave people feeling that they cannot control the important things in their lives. Without support they remain emotionally vulnerable to setbacks and upsets. For those who feel hopeless and helpless, depression is often the psychological result.'

4.6.81 Experiences of treatments

2 Psychological therapy

3 There was a strong feeling within the service user and carer topic group that the excerpt from
4 Howe (1995) in the section above highlights the reasons why many people opt for private
5 therapy; that is, that psychological treatment offered by the NHS in the form of CBT does not
6 go far enough in addressing the trauma experienced in childhood. The study by Ridge and
7 Ziebland (2006) confirms the opinions of the topic group and the testimony from the personal
8 accounts that people with 'deep and complex problems felt the need for longer term therapy'.
9 Those that have had long-term psychodynamic therapy report that it has been helpful in their
10 under- standing of themselves and their depression and that until they have worked through
11 and repaired the damage experienced in childhood, depression will be a major factor in the
12 person's life. The service user and carer topic group do acknowledge, however, that as there
13 has been little research into the efficacy of long-term psychodynamic therapy, it cannot be
14 recommended as a course of treatment in this guideline.

15 The study by Saver and colleagues (2007) points to the fact that few people received
16 information about psychological therapy and the different treatments, such as CBT and IPT.

17 Psychosocial interventions

18 This was a theme of both the personal accounts and the qualitative analysis. In the
19 qualitative analysis, people expressed a need for psychosocial interventions when they
20 attributed the cause of their depression to psychological processes rather than a 'chemical
21 imbalance' and to help them cope with negative thoughts.

22 Overall, people in the qualitative analysis were positive about counselling, as were people in
23 the personal accounts, although concerns were raised by two people (B and E). One found
24 counselling inadequate for her needs because it did not get to the 'root' of her depression
25 and indeed did not stop her depression from becoming more severe. Another felt that the
26 counselling she received was unsatisfactory: she was asked inappropriate questions,
27 incorrect assumptions were made about her life, and she felt that she did not talk enough
28 during the sessions. She felt that for counselling to be effective, the counsellor needed to
29 both listen and question skilfully.

30 In the qualitative analysis, people were generally positive about cognitive therapy, self-help
31 books and support groups, but less positive about relaxation therapy because people with
32 severe depression find it difficult to relax. The view of relaxation therapy is borne out in
33 personal account B. The personal accounts express mixed views about support groups: one
34 person (D) was very positive about them, but another (E) said that, while it was good to meet
35 other people, she gained no therapeutic value from attending.

36 Khan and colleagues (2007) synthesised qualitative studies of patient experiences of
37 depression management in primary care to develop a framework for a guided self-help
38 intervention.

39 Medication

40 There were mixed reports regarding medication. Some people did not find antidepressants
41 helpful, particularly in the form of a 'drug cocktail'; others were concerned about taking
42 tablets. In the literature, it emerged that taking medication could lead to ambivalent feelings:
43 on the one hand, people felt relief because medication helped them cope with difficulties in
44 their day-to-day life; on the other, they felt a lack of control. In the personal accounts, one
45 person (A) commented on the weight gain associated with the medication leading to self-
46 esteem issues and feeling more depressed. Others benefited from it; one person (B) felt
47 strongly that getting the appropriate medication promptly is vital and that there should be

intense support before the antidepressive effects are experienced. The majority of people in the qualitative analysis said that antidepressants were beneficial, despite some experiencing side effects.

Electroconvulsive therapy

This theme was only present in the qualitative analysis. The majority of people who had ECT had negative experiences, including loss of memory after treatment. Only one person had a positive experience with no side effects.

4.6.98 Coping strategies

It is evident from the personal accounts and the literature review that people who have had depression for a long time develop positive coping mechanisms that enable them to manage their illness. These mechanisms range from exercise (A) or personal faith (C), to readjusting one's life to be able to manage depression. The qualitative analysis also identified a number of coping strategies such as distraction, having a hobby, activities and voluntary work.

4.6.104 Employment

The theme of employment was only present in the personal accounts. To contextualise this theme, some of the literature regarding this topic that was not identified in the systematic search is briefly described below.

From the personal accounts there are issues for those with long-standing depression when it comes to accessing and remaining in employment. Several personal accounts spoke of difficulties in getting paid employment: one person (C) stated that both their college and job centre could not help until their condition was stable, and another (B) was self-employed when she became ill, was unable to work and had no income. In personal account G, the person had only worked in paid employment for 8 months in the 8 years he had had depression, but was doing voluntary work with mental health and disability organisations.

Other personal accounts spoke of experiences in work. One person (A) spoke of colleagues not being keen for her to return to work, and instead of returning to her normal activities she was marginalised from external meetings and confined to certain tasks. Another person (E) expressed the fear of getting too ill to work, but with the help of her GP did not have to say that the occasional day or week off with illness was because of depression. However, she also had the support of her manager in whom she confided and who helped with work pressures. In the qualitative analysis, some people commented that stressful situations at work contributed to the onset of their depression.

The issue of employment is also important to carers: in personal account H, the carer has built her career around self-employment so that she has time to care, but is also able to maintain a life outside caring.

Clinical research and government reports suggest that employment plays a part both in exacerbating stress leading to depression, but also, conversely, that it can be crucial component in aiding the recovery process. The Health and Safety Executive (2008) reported that in 2006/07, an estimated 530,000 people in the UK reported they were experiencing stress, depression or anxiety that was caused or exacerbated by their current or past employment. It was estimated that 13.8 million working days (full-day equivalent) were lost in 2006/07 through work-related stress, depression or anxiety. The Sainsbury Centre for Mental Health (2007) also identified the loss in productivity that occurs when employees come to work but function at less than full capacity because of ill health (termed 'presenteeism'). Fearing possible stigma or discrimination, people with mental health problems may turn up for work even if they are feeling unwell rather than be labelled as mentally ill by their employers and co-workers.

Once people with depression become too ill to work, they may remain absent from their place of employment or unemployed for considerable periods of time. The anecdotal evidence from the personal accounts suggests, however, that for people with depression a return to work or continuing with work can aid the recovery process. A report by Waddell and Burton (2006) concluded that work was generally beneficial for both physical and mental health and well-being. It advised that the type of employment should be healthy and safe, and should offer the individual some influence over how the work is done and a sense of self-worth. Overall, the beneficial effects of work were shown to outweigh the risks and to be much greater than the harmful effects of long- term unemployment or prolonged absence because of sickness.

A report by the Royal College of Psychiatrists (2008) found two studies that analysed employment schemes in people with mental health problems. In a systematic review of 11 RCTs comparing prevocational training or supported employment for people with severe mental illness with each other or with standard community care, Crowther and colleagues (2001) found that participants who received supported employment were more likely to be in competitive employment than those who received prevocational training (34% compared with 12% at 12 months). Rinaldi and colleagues (2007) examined a supported employment scheme run by South West London and St George's Mental Health NHS Trust. The results showed that, following the integration of employment specialists into CMHTs, there was a significant increase in the number of clients with various diagnoses (31% with depression – unspecified severity) engaged in mainstream work or educational activity at both 6 and 12 months. The conclusion drawn supports the use of individual placement specialists in clinical practice in CMHTs.

4.6.11 Recovery

In the study by Ridge and Ziebland (2006), the term 'recovery' is used to describe the process by which people learn to understand and then manage their illness. They explain that as the process of recovery develops, the person is able to assume responsibility for their illness through gaining insight into themselves, their thought processes, their concept of themselves and others around them, and their place in the world. Treatments and professionals were seen as the 'tools' needed to aid recovery. The term 'recovery' was the cause of significant debate in the service user and carer topic group and had different meanings for different people. For some it meant an absence of depressive symptoms and an ability to function fully to one's potential. But for other long-term sufferers, 'recovery' was a term that they would not use ('self- management' being perhaps a more appropriate term). For others the term 'recovery' was important in demonstrating the positive shift from being severely depressed with an inability to 'function normally', to perhaps currently living with dysthymia, where the user is able to live a full and productive life, with just a few residual symptoms that are manageable.

4.6.12 Families and carers

The literature search did not identify studies of carer experience and the two personal accounts offer very different perspectives, one from an adult caring for her partner (H) and one from a teenage boy caring for his mother (I). But several themes did emerge. The personal accounts both conveyed the experience that caring is rewarding but challenging. Both carers also spoke of the different aspects of caring: undertaking practical tasks for the person, and offering emotional support. Caring can radically change the relationship between partners and between parents and children. The carer in account H felt more like a mother than a partner and the young carer (I) said that he became an adult when he cared for his mother, but that she became a 'normal bossy Mum' again when she was well. Both carers reported that having interests that took them away from caring for a few hours was extremely important.

The needs of young carers should be recognised and addressed and recent publications from the Social Care Institute for Excellence and the Department of Health (Department of Health et al., 2008; Greene et al., 2008; Roberts et al., 2008; Department of Health et al., 2009) provide guidance on how this can be achieved. It should be recognised that young carers might marginalise themselves from their peer group and experience other social and educational disadvantage. The report by Roberts and colleagues (2008) suggests that the needs of young carers could be more effectively addressed by respecting their anxieties and acknowledging their input and skills. It is also recommended that young carers should be included in their family member's care planning.

The impact of depression on families and carers was a prolific theme in both the personal accounts and the qualitative analysis, with some people stating that depression was harder for family members and carers than for themselves. Some people remarked on the change of roles that occurred as a result of one person having depression. Many people also commented on the supportive nature of family members and carers, although some people had to cope with their depression alone.

4.7 Recommendations

Providing information and support, and obtaining informed consent

1. Make sure people with depression are aware of self-help groups, support groups and other local and national resources. [2004]

Advance decisions and statements

2. Consider developing advance decisions and advance statements collaboratively with people who have recurrent severe depression or depression with psychotic symptoms, and for those who have been treated under the Mental Health Act, in line with the Mental Capacity Act. Record the decisions and statements and include copies in the person's care plan in primary and secondary care, and give copies to the person and to their family or carer if the person agrees. [2009, amended 2017]

Supporting families and carers

3. When families or carers are involved in supporting a person with severe or chronic[a] depression, think about:

- providing written and verbal spoken information on depression and its management, including how families or carers can support the person
- offering a carer's assessment of their caring, physical and mental health needs if needed
- providing information about local family or carer support groups and voluntary organisations, and helping families or carers to access them
- discussing with the person and their family or carer about confidentiality and the sharing of information. [2009]

a Depression is described as 'chronic' if symptoms have been present more or less continuously for 2 years or more.

Working with people from diverse ethnic and cultural backgrounds

4. Be respectful of, and sensitive to, diverse cultural, ethnic and religious backgrounds when working with people with depression, and be aware of the possible variations in the presentation of depression these can cause. Ensure staff are competent in:

- culturally sensitive assessment
- using different explanatory models of depression
- addressing cultural and ethnic differences when developing and implementing treatment plans
- working with families from diverse ethnic and cultural backgrounds. [2009]

5. Provide all interventions in the preferred language of the person with depression if possible. [2004]

5 Organisation and delivery of services

5.1 Introduction

5.1.1 Current practice and aims of the review

Over the past 20 years, there has been a growing interest in the development of systems of care for managing depression. This work has been influenced by organisational developments in healthcare in the US, such as managed care and Health Maintenance Organisations (Katon et al. 1999), developments in the treatment of depression, the development of stepped care (Davison (2000)), and influences from physical healthcare (for example, chronic disease management (Wagner and Groves (2002)). A significant factor in driving these developments has been the recognition that for many people depression is a chronic and disabling disorder.

The implementation in the NHS of the various developments described in the introduction has been variable. Perhaps the model most widely adopted has been the stepped-care model within the IAPT programme (Department-of-Health (2007), but outside of demonstration sites and experimental studies (Layard 2006; Van Straten et al. 2006) there has not been a consistent adoption of any particular model of stepped care. Resource constraints have often been a significant limitation of these developments, but there have also been changes in mental healthcare policies that have influenced implementation, for example the varying developments of the attached professional role over the past 20 years (Bower and Sibbald 2000).

One consistent factor that links these developments is the limited evidence for most if not all of these interventions. The most notable exception is the evidence base for collaborative care, which has grown considerably in the past 20 years and has led some (for example, Simon 2006) to call for the widespread implementation of collaborative care. It should be noted that previous guidelines have heighted the presence of potentially important trial based research in this area (for example see systematic review by Gilbody and colleagues 2006) but that much of this evidence had previously been undertaken in the US and clear guidance could not be offered for UK primary care mental health services. In this updated guideline we have noted the conduct and publication of large scale trials and economic evaluations of collaborative care in the UK (Richards et al. 2013) and the present guideline incorporates new evidence with particular relevance to the UK.

Update 2017

5.1.2 Models of service delivery

There are a number of models of service delivery for people with depression which have featured in previous guidelines. In this guideline update, the over-arching term ‘enhanced care’ is used to refer to them all. This includes a number of interventions or models that often have some degree of overlap or where individual interventions are contained within more intensive or complex models. For example, collaborative care interventions (Gilbody et al. 2006) may include stepped care (Bower and Gilbody 2005) as a component (Katon et al. 1999, Unutzer et al. 2002), and also some element of medication management or brief psychological therapy. Some of the more prominent models are listed below.

The consultation-liaison model

This model (for example, Creed & Marks 1989, Darling & Tyler 1990, Gask et al.1997) is a variant of the training and education model (which is outside of the scope of the guideline), in that it seeks to improve the skills of primary care professionals and improve quality of care through improvements in their skills. However, rather than providing training interventions that teach skills in dealing with patients with depression in general, in this model specialists

enter into an ongoing educational relationship with the primary care team, in order to support them in caring for specific patients who are currently undergoing care. Referral to specialist care is only expected to be required in a small proportion of cases. A common implementation of this model involves a psychiatrist visiting practices regularly and discussing patients with primary care professionals.

The attached professional model

In this model (for example, Bower and Sibbald 2000), a mental health professional has direct responsibility for the care of a person (usually in primary care) focusing on the primary treatment of the problem/disorder, be it pharmacological or psychological. The co-ordination of care remains with the GP/primary care team. Contact is usually limited to treatment and involves little or no follow-up beyond that determined by the specific intervention offered (for example, booster sessions in CBT).

Stepped care

Stepped care (for example, Bower and Gilbody 2005) is a system for delivering and monitoring treatment with the explicit aim of providing the most effective yet least burdensome treatment to the patient first, and which has a self-correcting mechanism built in (that is, if a person does not benefit from an initial intervention they are 'stepped up' to a more complex intervention). Typically, stepped care starts by providing low-intensity interventions. In some stepped-care systems, low-intensity care is received by all individuals, although in other systems patients are stepped up to a higher intensity intervention on immediate contact with the service, for example if they are acutely suicidal (this later model is the one adopted in this guideline update and in the previous guideline).

Stratified (or matched care)

This is a hierarchical model of care (for example, Van Straten et al., 2006), moving from low- to high-intensity interventions, where at the patient's point of first contact with services they are matched to the level of need, and the consequent treatment is determined by the assessing professional in consultation with the patient.

Case management

This describes a system where an individual healthcare professional takes responsibility for the co-ordination of the care of an individual patient (for example, Gensichen et al. 2006), but is not necessarily directly involved in the provision of any intervention; it may also involve the co-ordination of follow-up.

Collaborative care

The collaborative care model (for example, Wagner 1997; Katon et al. 2001) emerged from the chronic disease model. A useful definition of the core elements of collaborative care have been provided by Gunn and colleagues (2006).

1. A multi-professional approach to patient care. This required that a general practitioner (GP) or family physician and at least one other health professional (for example, nurse, psychologist, psychiatrist, pharmacist) were involved with patient care.
2. A structured management plan. In line with introducing an organised approach to patient care 'systems' trials were required to offer practitioners access to evidence based management information. This could be in the form of guidelines or protocols. Interventions could include both pharmacological (for example, antidepressant medication) and non-pharmacological interventions (for example, patient screening, patient and provider education, counselling, cognitive behaviour therapy).

Update 2017

3. Scheduled patient follow-ups. A 'systems' approach required interventions to have an organised approach to patient follow-up. This is operationally-defined as one or more scheduled telephone or in-person follow-up appointments to provide specific interventions, facilitate treatment adherence, or monitor symptoms or adverse effects.
4. Enhanced inter-professional communication. This requires that the collaborative care intervention introduces mechanisms to facilitate communication between professionals caring for the depressed person. This can include team meetings, case-conferences, individual consultation/supervision, shared medical records, and patient-specific written or verbal feedback between care-givers.

In mental health services, collaborative care also typically includes a consultation liaison role with a specialist mental health professional and generic primary care staff.

Collaborative care may also include elements of many of the other interventions described above. In this guideline it is assumed that collaborative care, focused on the treatment and care of depression, is provided as part of a well-developed stepped care programme, and coordinated at either the primary or secondary care level. All sectors of care should be involved in order to ensure a comprehensive and integrated approach to mental and physical healthcare. Typically the programme of care is coordinated by a dedicated case manager supported by a multi-professional team. There will be joint determination with the service user regarding the care plan along with long-term coordination and follow-up.

5.1.3 Interventions included

The GC considered the range of interventions described above and the extent of current practice and decided to focus the reviews for this update on the following interventions: stepped care (including where possible matched care), collaborative care, the attached professional model and medication management. This was because they were the focus of considerable interest in the NHS and in the case of collaborative care considerable new evidence has emerged since the publication of the previous guideline. No additional studies were found for the attached professional models, so the GC decided that rather than performing a separate review they would comment on it, particularly in relation to collaborative care. The GC also decided to review medication management because there was evidence of increased use of this intervention in depression but considerable uncertainty as to whether the evidence supported medication management as a single intervention and not as part of a wider model of service delivery.

The increased focus on social inclusion and the role of employment in maintaining good mental health led the GC to also consider an updated review of employment but as no new studies were identified in the searches undertaken for this guideline the GC decided not to update the review undertaken for the previous guideline. For similar reasons the reviews of social support systems, crisis resolution and home treatment teams and day hospitals were not updated.

Definitions

The definitions adopted are as stated in section 5.1.1 with the exception of medication management, which is given below.

Medication management

Medication management (for example, Peveler et al., 1999) is an intervention aimed at improving patient adherence to medication. It is usually delivered by a pharmacist or nurse. It involves patient education about the nature and treatment of depression, the delivery of medication adherence strategies, the monitoring of side effects and the promotion of treatment adherence.

Update 2017

5.2 Review question

- For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 8. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 8: Clinical review protocol summary for the review of benefits and harms associated with different models for the coordination and delivery of services

Component	Description
Review question	For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services? (RQ1.1)
Population	Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by depression scale score for subthreshold and other groups For studies on relapse prevention: Adults whose depression has responded fully or partially to treatment according to DSM, ICD or similar criteria, or depressive symptoms as indicated by depression scale score
Intervention(s)	Models for the coordination and delivery of services • Collaborative care (simple and complex) • Medication management • Care co-ordination • Stepped care • Integrated care pathways (including primary care liaison or shared care)
Comparison	• Treatment as usual • Waitlist • Any alternative service delivery model
Critical outcomes	**Critical outcomes**: • Depression symptomology • Response • Remission • Relapse **Important but not critical outcomes:** • Service utilisation/resource use (e.g. antidepressant use)
Study design	RCTs and systematic reviews

Update 2017

5.2.1 Clinical evidence

The GC selected an existing, high-quality systematic review as the main source of RCTs for this review (Coventry et al. 2014; 78 RCTs). Additional RCTs were identified from the previous iteration of the NICE guideline (12 RCTs), through another systematic review identified during the search process (van Straten 2015; 5 RCTs), through our own update searches including those conducted for other review questions (45 RCTs) and via handsearch (6 RCTs). In total 146 RCTs were assessed for eligibility at full text and 73 were included. Following inclusion each RCT (or study arm, in the case of multiple-arm RCTs) was categorised by format of service delivery using the checklist set out within the review protocol for this question (Appendix F). The categories were collaborative care (simple: 42 [relapse

prevention: 1], complex: 11); stepped care (total: 4, relapse prevention: 1); medication management (10); care coordination (4); integrated care pathways (primary care liaison: 2, integrated pathways: 1). Each of these reviews is presented below; relapse prevention delivery models are presented together irrespective of category.

5.2.1.15 Collaborative care

53 RCTs were categorised as collaborative care and included in this review: Adler, Bungay et al. (2004), Aragonès, Piñol et al. (2012), Araya, Rojas et al. (2003), Berghöfer, Hartwich et al. (2012), Bosanquet, Adamson et al. (2017), Bruce, Ten Have et al. (2004), Buszewicz, Griffin et al. (2010), Capoccia, Boudreau et al. (2004), Chen, Conwell et al. (2015), Chew-Graham, Lovell et al. (2007), Ciechanowski, Wagner et al. (2004), Cole, McCusker et al. (2006), Cooper, Ghods Dinoso et al. (2013), Datto, Thompson et al. (2003), Dietrich, Oxman et al. (2004), Dwight-Johnson, Aisenberg et al. (2011), Ell, Unützer et al. (2007), Finley, Rens et al. (2003), Fortney, Pyne et al. (2007), Fortney, Pyne et al. (2013), Gensichen, vonKorff et al. (2009), Hedrick, Chaney et al. (2003), Huijbregts, Jong et al. (2013), Katon (1996a), Katon (1996b), Katon, Von Korff et al. (1999), Katon, Von Korff et al. (2001), Katzelnick, Simon et al. (2000), Lewis, Adamson et al. (2016), Ludman (2007a), Ludman (2007b), Ludman (2007c), McCusker, Cole et al. (2008), Melville, Reed et al. (2014), Menchetti, Sighinolfi et al. (2013), Oslin, Sayers et al. (2003), Patel, Weiss et al. (2010), Richards, Lovell et al. (2008), Richards, Hill et al. (2013), Ross, TenHave et al. (2008), Rost, Nutting et al. (2001), Rost, Nutting et al. (2002), Rubenstein, Parker et al. (2002), Simon (2000a), Simon (2000b), Simon (2004a), Simon (2004b), Simon, Ralston et al. (2011), Unutzer, Katon et al. (2002), Vlasveld, Feltz-Cornelis et al. (2012), Wells (2000a), Wells (2000b), Yeung, Shyu et al. (2010).

These 53 RCTs were separated into 3 different comparisons; simple collaborative care versus control, complex collaborative care versus control and head-to-head comparisons of different forms of collaborative care.

An overview of the trials included in the meta-analyses can be found in Table 9 and Table 10. The majority of the data is from US studies conducted in primary care settings in white, female populations in their mid-40s. Further information about both included and excluded studies can be found in Appendix J1.1.

Summary of findings can be found in Table 11 and Table 12. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

Data were available for all critical and important outcomes.

Table 9: Study information table for trials included in the meta-analysis of collaborative care compared to control

	Simple collaborative care versus control	Complex collaborative care versus control
Total no. of studies (N[1])	40 (12,705)	11 (3,829)
Study ID	Adler 2004[2] Aragones 2012[3] Araya 2003[4] Berghofer 2012[5] Bosanquet submitted[6] Bruce 2004[7] Buszewicz 2011[8] Capoccia 2004[9] Chen 2015[10] Chew-Graham 2007[11]	Ciechanowski 2004[43] Ell 2007[44] Fortney 2007[45] Hedrick 2003[46] Huijbregts 2013[47] Katon 1996a[48] Katon 1996b [49] Melville 2014[50] Simon 2004b[51] Unutzer 2002[52]

Update 2017

	Simple collaborative care versus control	Complex collaborative care versus control
	Cole 2006[12] Datto 2003[14] Dietrich 2004[15] Dwight-Johnson 2010[16] Finley 2003[17] Gensichen 2009[18] Katon 1999[19] Katon 2001[20] Katzelnick 2000[21] Lewis 2016[22] Ludman 2007a[23] Ludman 2007b[24] Ludman 2007c[25] McCusker 2008[26] Menchetti 2013[27] Oslin 2003[28] Patel 2010[29] Richards 2008[30] Richards 2013[31] Ross 2008[32] Rost 2001[33] Rost 2002[34] Rubenstein 2002[35] Simon 2000a[36] Simon 2000b[37] Simon 2004a[38] Simon 2011[39] Wells 2000a[40] Wells 2000b[41] Yeung 2010[42]	Vlasveld 2012[53]
Country	USA[2,7,9,14,15,16,17,19,20,21,23,24,25,27,28,32,33,34,35,36,37,38,39,40,41,42] Spain[3] Chile[4] UK[6,8,11,22,30,31] China[10] Canada[12,26] Germany[5,18] India[29]	USA
Age (mean)	NR[10] <40[9,16] 40-50[2,3,4,5,8,12,14,15,18,19,20,21,23,24,25,27,29,30,31,33,34,35,36,37,38,39,40,41,42] 51-64[7,17,28,32] >=65[6,11,22,26]	NR[47,53] 40-64[45,46,48,49,51] >=65[43,44,52]
Sex	>50% male[28,32] >50% female[2,3,4,5,6,7,8,9,10,11,12,14,15,16,17,18,19,20,21,22,23,24,25,26,27,29,30,31,33,34,35,36,37,38,39,40,41,42]	NR[47,53] >50% male[45,46] >50% female[43,44,48,49,51,52]
Ethnicity	NR[3,5,10,11,16,17,18,26,27,29,33,36,37]	NR[44,47,53]

Update 2017

	Simple collaborative care versus control	Complex collaborative care versus control
	>50% white[2,6,7,8,9,12,14,15,19,20,21,22,23,24,25,30,31,34,35,38,39,40,41] >50% non-white[4,28,32,42]	>50% white[43,45,46,51,52]
Treatment setting	Primary care	Primary care
Intervention	Simple collaborative care	Complex collaborative care
Comparison	Care as usual	Care as usual
Notes: [1] Number randomised, Adler 2004[2], Aragones 2012[3], Araya 2003[4], Berghofer 2012[5], Bosanquet submitted[6], Bruce 2004[7], Buszewicz 2011[8], Capoccia 2004[9], Chen 2015[10], Chew-Graham 2007[11], Cole 2006[12], Datto 2003[14], Dietrich 2004[15], Dwight-Johnson 2010[16], Finley 2003[17], Gensichen 2009[18], Katon 1999[19], Katon 2001[20], Katzelnick 2000[21], Lewis 2016[22], Ludman 2007a[23], Ludman 2007b[24], Ludman 2007c[25], McCusker 2008[26], Menchetti 2013[27], Oslin 2003[28], Patel 2010[29], Richards 2008[30], Richards 2013[31], Ross 2008[32], Rost 2001[33], Rost 2002[34], Rubenstein 2002[35], Simon 2000a[36], Simon 2000b[37], Simon 2004a[38], Simon 2011[39], Wells 2000a[40], Wells 2000b[41], Yeung 2010[42], Ciechanowski 2004[43], Ell 2007[44], Fortney 2007[45], Hedrick 2003[46], Huijbregts 2013[47], Katon 1996a[48], Katon 1996b [49], Melville 2014[50], Simon 2004b[51], Unutzer 2002[52], Vlasveld 2012[53]		

Table 10: Study information table for trials included in the meta-analysis of collaborative care compared to active intervention

	Collaborative care versus active intervention
Total no. of studies (N[1])	2 (496)
Study ID	Cooper 2013[2] Fortney 2013[3]
Country	USA
Baseline depression symptoms	CES-D: 29.84[2] Hopkins Symptom Checklist: 1.9[3]
Age (mean)	46.5[2] 47.2[3]
Sex (% female)	77%[2] 81%[3]
Ethnicity (% white)	NR
Treatment setting	Primary care
Intervention	Standard Collaborative Care[2] Telemedicine Based Collaborative Care: stepped care, provided via telephone or video-conference dependent upon severity[3]
Comparison	Patient-centred collaborative care: as in the standard condition, but access barriers were also explored[2] Practice Based Collaborative Care: watchful waiting or antidepressant treatment provided[3]
Notes: [1] Number randomised [2]Cooper 2013; [3]Fortney 2013	

Update 2017

Table 11: Summary of findings table for the comparison of collaborative care versus control

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Collaborative care (95% CI)
Depression symptoms- 6 months	10,602 (47 studies) 6 months	⊕⊖⊖⊖ **very low**[3,1] due to risk of bias, inconsistency			The mean depression symptoms- 6 months in the intervention groups was **0.31 lower** (0.39 to 0.23 lower)
Depression symptoms- Simple collaborative care	7,881 (36 studies) 6 months	⊕⊖⊖⊖ **very low**[3,1] due to risk of bias, inconsistency			The mean depression symptoms- simple collaborative care in the intervention groups was **0.32 lower** (0.41 to 0.22 lower)
Depression symptoms- Complex collaborative care	3,079 (11 studies) 6 months	⊕⊖⊖⊖ very low[3,1] due to risk of bias, inconsistency			The mean depression symptoms- complex collaborative care in the intervention groups was **0.28 lower** (0.43 to 0.13 lower)
Depression symptoms at follow-up	4539 (9 studies) 12 months	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, inconsistency			The mean depression symptoms at follow-up in the intervention groups was **0.23 lower** (0.4 to 0.07 lower)
Depression symptoms at follow-up - Simple collaborative care	2568 (6 studies) 12 months	⊕⊕⊖⊖ **low**[1] due to risk of bias			The mean depression symptoms at follow-up - simple collaborative care in the intervention groups was **0.21 lower** (0.3 to 0.12 lower)
Depression symptoms at follow-up - Complex collaborative care	1971 (3 studies) 12 months	⊕⊖⊖⊖ **very low**[1,2,3] due to risk of bias, inconsistency, imprecision			The mean depression symptoms at follow-up - complex collaborative care in the intervention groups was **0.27 lower** (0.72 lower to 0.17 higher)
Non-response at follow-up	3278 (10 studies) 12 months	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, inconsistency	**RR 0.72** (0.63 to 0.81)	**Study population**	
				748 per 1000	**209 fewer per 1000** (from 142 fewer to 277 fewer)
				Moderate	

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Collaborative care (95% CI)
				681 per 1000	**191 fewer per 1000** (from 129 fewer to 252 fewer)
Non-response at follow-up- Simple collaborative care	895 (4 studies) 12 months	⊕⊖⊖⊖ **very low**[1,3,4] due to risk of bias, inconsistency, imprecision	**RR 0.66** (0.47 to 0.92)	**Study population**	
				598 per 1000	**203 fewer per 1000** (from 48 fewer to 317 fewer)
				Moderate	
				394 per 1000	**134 fewer per 1000** (from 32 fewer to 209 fewer)
Non-response at follow-up - Complex collaborative care	2383 (6 studies) 12 months	⊕⊕⊖⊖ **low**[1] due to risk of bias	**RR 0.75** (0.66 to 0.85)	**Study population**	
				802 per 1000	**201 fewer per 1000** (from 120 fewer to 273 fewer)
				Moderate	
				750 per 1000	**188 fewer per 1000** (from 112 fewer to 255 fewer)
Antidepressant use- 6 months	0 (31 studies) 6 months	⊕⊖⊖⊖ **very low**[3,1] due to risk of bias, inconsistency	**RR 1.39** (1.26 to 1.52)	**Study population**	
				See comment	-
				Moderate	
				0 per 1000	-
Antidepressant use- 6 months - Simple collaborative care	0 (22 studies)	⊕⊖⊖⊖ **very low**[3,1] due to risk of bias, inconsistency	**RR 1.45** (1.26 to 1.66)	**Study population**	
				See comment	-
				Moderate	
				0 per 1000	-
Antidepressant use- 6 months - Complex collaborative care	0 (10 studies)	⊕⊖⊖⊖ **very low**[3,4] due to risk of bias, imprecision	**RR 1.29** (1.2 to 1.38)	**Study population**	
				See comment	-

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Collaborative care (95% CI)
				Moderate	
				0 per 1000	-
Antidepressant use at follow-up	3618 (10 studies) 12 months	⨁⊖⊖⊖ **very low**[1,3,4] due to risk of bias, inconsistency, imprecision	**RR 1.18** (1.03 to 1.35)	**Study population**	
				534 per 1000	**96 more per 1000** (from 16 more to 187 more)
				Moderate	
				550 per 1000	**99 more per 1000** (from 16 more to 193 more)
Antidepressant use at follow-up - Simple collaborative care	1383 (6 studies) 12 months	⨁⊖⊖⊖ **very low**[1,3,4] due to risk of bias, inconsistency, imprecision	**RR 1.14** (0.9 to 1.46)	**Study population**	
				485 per 1000	**68 more per 1000** (from 48 fewer to 223 more)
				Moderate	
				380 per 1000	**53 more per 1000** (from 38 fewer to 175 more)
Antidepressant use at follow-up - Complex collaborative care	2235 (4 studies) 12 months	⨁⊖⊖⊖ **very low**[1,3] due to risk of bias, imprecision	**RR 1.26** (1.17 to 1.35)	**Study population**	
				565 per 1000	**147 more per 1000** (from 96 more to 198 more)
				Moderate	
				619 per 1000	**161 more per 1000** (from 105 more to 217 more)
Non-remission at 6 months (simple collaborative care)	211 (1 study)	⨁⊖⊖⊖ **very low**[1,3,5] due to risk of bias, imprecision	**RR 0.81** (0.66 to 1)	**688 per 1000**	**131 fewer per 1000** (from 234 fewer to 0 more)
Non-remission at follow-up	395 (2 studies) 12 months	⨁⊖⊖⊖ **very low**[2,3,6] due to risk of bias, inconsistency, imprecision	**RR 0.58** (0.38 to 0.89)	**788 per 1000**	**331 fewer per 1000** (from 87 fewer to 488 fewer)

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Collaborative care (95% CI)
Non-remission at follow-up - simple collaborative care	214 (1 study) 12 months	⊕⊕⊖⊖ **low**[6,7] due to risk of bias, imprecision	**RR 0.47** (0.37 to 0.59)	**913 per 1000**	**484 fewer per 1000** (from 375 fewer to 575 fewer)
Non-remission at follow-up - complex collaborative care	1041 (1 study) 12 months	⊕⊕⊖⊖ **low**[3,6] due to risk of bias, imprecision	**RR 0.73** (0.56 to 0.95)	**64 per 1000**	**17 fewer per 1000** (from 3 fewer to 28 fewer)

Notes:
[1] ROB high or unclear across multiple domains in most studies
[2] I-squared >80%
[3] 95% CI crosses one clinical decision threshold
[4] I-squared >50%
[5] ROB high or unclear across multiple domains
[6] ROB high or unclear across a two to three domains
[7] OIS not met (<300 events)

Table 12: Summary of findings table for the comparison of collaborative care versus other active comparison

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with patient centred/practice-based collaborative care	Risk difference with standard/telebased collaborative care (95% CI)
Simple collaborative care: Standards CC vs patient centred CC- remission at follow-up	132 (1 study) 12 months	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision	**RR 1.27** (0.81 to 1.98)	**Study population**	
				328 per 1000	**89 more per 1000** (from 62 fewer to 322 more)
				Moderate	
				328 per 1000	**89 more per 1000** (from 62 fewer to 321 more)
Telebased CC vs Practice based CC- response- 6 months	318 (1 study) 6 months	⊕⊕⊕⊖ **moderate** due to risk of bias	**RR 3.02** (2.02 to 4.51)	**Study population**	
				152 per 1000	**306 more per 1000** (from 155 more to 532 more)
				Moderate	
				152 per 1000	**307 more per 1000** (from 155 more to 534 more)
				Study population	

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with patient centred/practice-based collaborative care	Risk difference with standard/telebased collaborative care (95% CI)
Telebased CC vs practice based CC- response at follow-up	287 (1 study) 12 months	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision	**RR 2.54** (1.79 to 3.61)	**208 per 1000**	**320 more per 1000** (from 164 more to 543 more)
				Moderate	
				208 per 1000	**320 more per 1000** (from 164 more to 543 more)

Notes
[1] ROB high or unclear across two to three domains
[2] 95% CI crosses one clinical decision threshold

5.2.1.1.1 Collaborative care: subgroup analysis

The collaborative care dataset was large enough to allow for subgroup analysis to further examine the results. The GC were particularly interested in examining whether collaborative care was more or less effective in older adults, in BME groups or in people with chronic depression, and whether case manager background, whether or not a psychological intervention was provided, the number of contacts provided as part of the intervention and whether a stepped care algorithm was used affected the utility of collaborative care.

In older adults collaborative care overall had a small beneficial effect on depressive symptoms at 6 month follow-up (SMD=-0.45 [-0.71,-0.18]), with this effect being clearer within the simple (the larger dataset) than the complex group (SMD simple=-0.48 [-0.77, -0.19] versus complex=-0.34 [-1.25, 0.58]). In BME patients collaborative care had a small-moderate beneficial effect on depressive symptoms at 6 month follow-up (SMD=-0.48 [-0.87,-0.09]). The beneficial effect was much smaller in patients with chronic depression (SMD=-0.23 [-0.35, -0.10]).

The professional background of the case manager did not impact upon the effectiveness of the intervention as measured by depressive symptoms (SMD mental health background=-0.31 [-0.40, -0.23] versus non-mental health background=-0.30 [-0.47, -0.13]). A greater number of contacts did appear to increase the effect size, with a small-moderate effect in those who received over 13 contacts (SMD=-0.40 [-0.69, -0.11]) compared with those who received less than 13 sessions (SMD=-0.29 [-0.36, -0.22]). The inclusion of a psychological intervention component within the collaborative care intervention did not make a significant difference to effectiveness as measured by depressive symptoms at endpoint (SMD psychological intervention=-0.33 [-0.42, -0.24] compared with non-psychological intervention =-0.28 [-0.44, -0.12]). Collaborative care that included a stepped care algorithm was most effective (SMD=-0.46 [-0.68, -0.25]), followed by medication algorithm (SMD=-0.31 [-0.41, -0.20]), decision support (SMD=-0.30 [-0.52, -0.08]), and finally no stepped care component (SMD=-0.24 [-0.31, -0.18]).

5.2.1.2 Stepped care

3 RCTs were categorised as stepped care and included in this review: Bauer, Pretorius et al. (2009), Oladeji, Kola et al. (2015), Van't Veer-Tazelaar, Smit et al. (2010).

An overview of the trials included in the meta-analyses can be found in Table 13. Further information about both included and excluded studies can be found in Appendix J1.1.

Update 2017
Update 2017

Summary of findings can be found in Table 14. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

No data were available for the critical outcome of response.

Table 13: Study information table for trials included in the meta-analysis of stepped care compared with control

	Stepped care versus control
Total no. of studies (N[1])	3 (552)
Study ID	Bauer 2009[2] Oladeji 2015[3] van't Veer Tazelaar 2009[4]
Country	Germany[2] Nigeria[3] Netherlands[4]
Baseline depression symptoms	NR[2] PHQ-9=11.3 (3.61)[3] CES-D=21.6 (5.1)[4]
Age (mean)	48.2[2] 43.2[3] 81.4[4]
Sex (% female)	60%[2] 80%[3] 74%[4]
Ethnicity (% white)	98%[2] NR[3,4]
Treatment setting	Inpatient[2] Primary care[3,4]
Intervention	Standardised stepwise drug treatment regime (SSTR)[2] Stepped care, dependent upon the PHQ-9 score; 24 weeks[3] Stepped care: step 1; watchful waiting, step 2: Cognitive behaviour therapy–based bibliotherapy, step 3: Brief cognitive behaviour therapy–based problem solving, step 4: referral to primary care ; 52 weeks[4]
Comparison	Care as usual
Notes: [1] Number randomised [2]Bauer 2009; [3]Oladeji 2015; [4]van't Veer Tazelaar 2009	

Table 14: Summary of findings table for the comparison of stepped care versus control

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Stepped care (95% CI)
Remission at endpoint	148 (1 study)	⨁⨁⊖⊖ **low**[1,2] due to risk of bias, imprecision	**RR 1.38** (0.97 to 1.96)	**Study population**	
				392 per 1000	**149 more per 1000** (from 12 fewer to 376 more)
				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Stepped care (95% CI)
				392 per 1000	**149 more per 1000** (from 12 fewer to 376 more)
Depression symptoms at endpoint PHQ-9	201 (1 study)	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision		The mean depression symptoms at endpoint in the control groups was **5.5**	The mean depression symptoms at endpoint in the intervention groups was **1.4 lower** (2.87 lower to 0.07 higher)
Antidepressant use	170 (1 study) 6 months	⊕⊖⊖⊖ **very low**[3,4] due to risk of bias, imprecision	**RR 1.19** (0.75 to 1.89)	**274 per 1000**	**52 more per 1000** (from 68 fewer to 244 more)

Notes:
[1] ROB high or unclear in two to three domains
[2] 95% CI crosses one clinical decision threshold
[3] High or unclear ROB in most domains
[4] 95% CI crosses two clinical decision thresholds

Update 2017

5.2.1.31 Medication management

10 RCTs were categorised as medication management and included in this review: Brook, van Hout et al. (2005), Katon (1995a), Katon (1995b), Lobello, Reddy et al. (2010), Ludman (2007a), Perahia, Quail et al. (2008), Peveler, George et al. (1999), Rickles, Svarstad et al. (2005), Rubio-Valera, Bosmans et al. (2013), Swindle, Rao et al. (2003).

An overview of the trials included in the meta-analyses can be found in Table 15. Further information about both included and excluded studies can be found in Appendix J1.1.

Summary of findings can be found in Table 16. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

No data were available for the critical outcomes of response and remission.

Table 15: Study information table for trials included in the meta-analysis of medication management compared with control

	Medication management versus control
Total no. of studies (N[1])	10 (2609)
Study ID	Brook 2005[2] Katon 1995a[3] Katon 1995b[4] Lobello 2010[5] Ludman 2007a[6] Perahia 2008[7] Peveler 1999[8] Rickles 2005[9]

	Medication management versus control
	Rubio-Valera 2013[10] Swindle 2003[11]
Country	Netherlands[2] 11 European countries[7] UK[8] USA[3,4,5,6,9,11] Spain[10]
Baseline depression symptoms	SCL-13= 2.94 (0.62)[2], NR[3,4,5,6,11] HAMD-17: Intervention=21.6 (4.0); Control=21.7 (4.2)[7], HADS= 12.6 (4.4)[8], BDI-II: PGEM 28.9 (8.15); UC 27.0 (8.40)[9], PHQ-9= 15.9[10]
Age (mean)	42.4 (8.9)[2], 51.1[3], 42.8[4], 44.5[5], 50.2[6], 46 (13)[7], 45.3 (21-83)[8], 38 (12)[9], 46.6[10], 56.2[11]
Sex (% female)	71.0%[2], 72.0%[3], 82.0%[4], 73.0%[5], 69.0%[6], 64.0%[7], 74.0%[8], 84.0%[9], 75.4%[10], 3.0%[11]
Ethnicity (% white)	NR[2,3,4,8,9,10], 87.3%[5], 86.0%[6], 99%[7], 85.5[11]
Treatment setting	Primary care[2,8,10] NR[3,4,5,6,11] Outpatients[7] Pharmacies[9]
Intervention	Pharmacy-based coaching: 3x 10-20 min sessions of one to one coaching about their medication use, and received a take-home video to improve their knowledge[2] Medication managment[3,4,5,6,11] Telephone Care Management: 3x telephone sessions over 12 weeks[7] Medication counselling: 2x sessions delivered by a nurse[8] Pharmacist Guided Education and Monitoring (PGEM): 3 monthly telephone calls, medication management and education[9] Community pharmacist intervention[10]
Comparison	Care as usual[2,3,4,5,6,9,10,11] Treatment as usual: duloxetine 60-120mg/day[7] Care as usual: leaflet provided[8]
Notes: [1]Number randomised [2]Brook 2005, [3]Katon 1995a, [4]Katon 1995b, [5]Lobello 2010, [6]Ludman 2007a, [7]Perahia 2008, [8]Peveler 1999, [9]Rickles 2005, [10]Rubio-Valera 2013, [11]Swindle 2003	

Update 2017

Table 16: Summary of findings table for the comparison of medication management versus control

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Medication management (95% CI)
Mean change in depression scores (SMD) **BDI/PHQ-9/HADS/HAM-D/SCL-20**	0 (9 studies)	⊕⊖⊖⊖ **very low**[1,3] due to risk of bias, inconsistency			The mean change in depression scores in the intervention groups was **0.13 less** (0.33 less to 0.06 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Medication management (95% CI)
Mean change in depression scores at follow-up (MD) **BDI**	219 (1 study) 12 months	⨁⊖⊖⊖ **very low**[1,2,3] due to risk of bias, imprecision		The mean change in depression scores at follow-up in the control groups was **19.9**	The mean change in depression scores at follow-up in the intervention groups was **2 lower** (4.86 lower to 0.86 higher)
Antidepressant use at endpoint	0 (4 studies)	⨁⊖⊖⊖ **very low**[1,2,3] due to risk of bias, inconsistency, imprecision	Not estimable	See comment	-

Notes:
[1] ROB high or unclear across two to three domains
[2] OIS not met (<400 participants)
[3] 95% CI crosses two clinical decision thresholds

5.2.1.41 Care co-ordination

2 4 RCTs were categorised as care co-ordination and included in this review: Landis, Gaynes
3 et al. (2007), Mann, Blizard et al. (1998)[trial 2], McMahon, Foran et al. (2007), Uebelacker,
4 Marootian et al. (2011).

5 An overview of the trials included in the meta-analyses can be found in Table 17. Further
6 information about both included and excluded studies can be found in Appendix J1.1.

7 Summary of findings can be found in Table 18. The full GRADE evidence profiles and
8 associated forest plots can be found in Appendices L and M.

9 No data were available for the critical outcomes of response and remission.

10 **Table 17: Study information table for trials included in the meta-analysis of care co-**
11 **ordination compared with control**

	Care co-ordination versus control
Total no. of studies (N[1])	4 (722)
Study ID	Landis 2007[2] Mann 1998[3] McMahon 2007[4] Uebelacker 2011[5]
Country	USA[2,5] UK[3,4]
Diagnosis	Depression
Baseline depression symptoms	NR
Age (mean)	39.7[2] 44.2[3]

	Care co-ordination versus control
	NR[4] 39.1[5]
Sex (% female)	96.0%[2] 78.0%[3] NR[4] 95.0%[5]
Ethnicity (% white)	62.2%[2] NR[3,4] 0%[5]
Treatment setting	NR
Intervention	Care coordination
Comparison	Care as usual
Notes: [1] Number randomised [2]Landis 2007, [3]Mann 1998, [4]McMahon 2007, [5]Uebelacker 2011	

Table 18: Summary of findings table for the comparison of care co-ordination versus control

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Care co-ordination (95% CI)
Mean change in depression scores at endpoint	0 (4 studies)	⨁⨁⊖⊖ **low**[1] due to risk of bias			The mean change in depression scores at endpoint in the intervention groups was **0.05 standard deviations lower** (0.35 lower to 0.25 higher)
Antidepressant adherence at follow-up	0 (3 studies) 12 months	⨁⊖⊖⊖ **very low**[2,3,4] due to risk of bias, inconsistency, imprecision	**RR 1.79** (0.68 to 4.72)	**Study population**	
				See comment	-
				Moderate	
				0 per 1000	-

Notes:
[1] ROB high or unclear in two to three domains
[2] ROB high or unclear across multiple domains
[3] I-squared>50%
[4] 95% CI crosses two clinical decision thresholds

5.2.1.53 Integrated care pathways

3 RCTs were categorised as integrated care pathways and included in this review:Blanchard, Waterreus et al. (1995), Dobscha, Corson et al. (2006), Krahn, Bartels et al. (2006).

Within this the Dobscha 2006 and Blanchard 1995 studies examined primary care liaison and the Krahn 2006 study looked at integrated care pathways.

An overview of the trials included in the meta-analyses can be found in Table 19. Further information about both included and excluded studies can be found in Appendix J1.1.

Summary of findings can be found in Table 20. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

No data were available for the critical outcomes of response and remission.

Table 19: Study information table for trials included in the meta-analysis of integrated care compared with control

	Integrated care versus control
Total no. of studies (N[1])	3 (2002)
Study ID	Blanchard 1995[2] Dobscha 2006[3] Krahn 2006[4]
Country	UK[2] USA[3,4]
Diagnosis	Depression[2,4] Depressive symptoms[3]
Baseline depression symptoms	NR[2] SCL-20: 1.9[3] CES-D: 24.95[4]
Age (mean)	76.3[2] 57.0[3] 73.9 (6.6)[4]
Sex (% female)	85.0%[2] 6.9%[3] 30.7%[4]
Ethnicity (% white)	NR[2] 47%[3] 45.1%[4]
Treatment setting	NR[2] Primary care[3,4]
Intervention	Integrated care[2] Primary care liaison; decision support programme[3] Integrated care: mental health and substance abuse services co-located in primary care[4]
Comparison	Care as usual[2,3] Enhanced care as usual: referrals to specialty providers within 2-4 weeks[4]
Notes: [1] Number randomised [2]Blanchard 1995, [3]Dobscha 2006, [4]Krahn 2006	

Table 20: Summary of findings table for the comparison of integrated care versus control

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Integrated care (95% CI)
Mean change in depression scores at endpoint	0 (3 studies)	⊕⊖⊖⊖ **very low**[4,2] due to risk of bias, inconsistency			The mean change in depression scores at endpoint in the intervention groups was **0.05 standard deviations lower** (0.26 lower to 0.16 higher)
Mean change in depression scores at endpoint - Integrated care vs control	0 (2 studies)	⊕⊖⊖⊖ **very low**[4,2,5] due to risk of bias, inconsistency, imprecision			The mean change in depression scores at endpoint - integrated care vs control in the intervention groups was **0.19 standard deviations lower** (0.55 lower to 0.17 higher)
Mean change in depression scores at endpoint - Integrated care vs speciality referral system	0 (1 study)	⊕⊕⊖⊖ **low**[1] due to risk of bias			The mean change in depression scores at endpoint - integrated care vs speciality referral system in the intervention groups was **0.08 standard deviations higher** (0.03 lower to 0.19 higher)
Mean change in depression scores at follow-up	375 (1 study) 12 months	⊕⊕⊖⊖ **low**[2,3] due to risk of bias, imprecision			The mean change in depression scores at follow-up in the intervention groups was **0.01 higher** (0.11 lower to 0.13 higher)
Antidepressant adherence	0 (2 studies)	⊕⊖⊖⊖ **very low**[4,2,3] due to risk of bias, inconsistency, imprecision	Not estimable	See comment	-

Notes:
[1] ROB high or unclear in multiple domains
[2] ROB high or unclear in two to three domains
[3] OIS not met (<400 participants)

5.2.1.61 Relapse prevention

2 RCTs were categorised as relapse prevention and included in this review: Apil, Hoencamp et al. (2012), Katon, Von Korff et al. (2001).

Within this the Apil 2012 study examined stepped care and the Katon 2001 study looked at collaborative care.

An overview of the trials included in the meta-analyses can be found in Table 21. Further information about both included and excluded studies can be found in Appendix J1.1.

Summary of findings can be found in Table 22. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

No data were available for the critical outcomes of response and remission.

Table 21: Study information table for trials included in the meta-analysis of relapse prevention interventions compared with control

	Relapse prevention interventions versus control
Total no. of studies (N[1])	2 (486)
Study ID	Apil 2012[2] Katon 2001
Country	Netherlands[2] USA[3]
Diagnosis	Depression[2] Subthreshold symptoms[3]
Baseline depression symptoms	CES-D: 17.2[2] NR[3]
Age (mean)	65.6 (8.3)[2] 46.0[3]
Sex (% female)	72.1%[2] 74.0%[3]
Ethnicity (% white)	NR[2] 90.2[3]
Treatment setting	Outpatients[2] NR[3]
Intervention	Stepped care: step 1: watchful waiting, step 2: nurse contacted participants to ensure treatment adherence every 2 weeks, step 3: 12x 45 min weekly sessions of coping with depression course, step 4: referred for specialist mental healthcare from physician or psychotherapist[2] Collaborative care[3]
Comparison	Care as usual
Notes: [1] Number randomised [2]Apil 2012, [3]Katon 2001	

Table 22: Summary of findings table for the comparison of relapse prevention interventions versus control

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Relapse prevention (95% CI)
Collaborative care (simple)- depression symptoms at endpoint	327 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision		The mean collaborative care (simple)- depression symptoms at endpoint in the control groups was **0.73**	The mean collaborative care (simple)- depression symptoms at endpoint in the intervention groups was **0.09 lower** (0.2 lower to 0.02 higher)
Collaborative care (simple)- relapse at follow-up	386 (1 study) 12 months	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	**RR 1.01** (0.77 to 1.33)	**Study population**	
				345 per 1000	**3 more per 1000** (from 79 fewer to 114 more)
				Moderate	
				345 per 1000	**3 more per 1000** (from 79 fewer to 114 more)
Stepped care - relapse at follow-up	136 (1 study) 12 months	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 1.26** (0.74 to 2.15)	**Study population**	
				258 per 1000	**67 more per 1000** (from 67 fewer to 297 more)
				Moderate	
				258 per 1000	**67 more per 1000** (from 67 fewer to 297 more)

Notes:
[1] ROB high or unclear in multiple domains
[2] OIS not met (<400 participants)
[3] 95% CI crosses one clinical decision threshold
[4] 95% CI crosses two clinical decision thresholds

5.2.23 Economic evidence

The systematic search of the literature identified 13 studies on the cost effectiveness of different models for the coordination and delivery of services for adults with depression. Details on the methods used for the systematic search of the economic literature, including inclusion criteria for each review question, are described in Chapter 3. Full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix Q. Completed methodology checklists of the studies are provided in Appendix P. 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 R.

Update

5.2.2.11 Collaborative care

The systematic search of the literature identified 3 UK economic studies on simple collaborative care (Bosanquet et al., 2017; Green et al., 2014; Lewis et al., 2017) and only one UK economic study on complex collaborative care (Morriss et al., 2016); following the hierarchy of inclusion criteria regarding country settings, 2 Dutch studies assessing the cost effectiveness of complex collaborative care were also included in the review (Goorden 2014 and 2015). In addition, the search identified one US study assessing the cost effectiveness of simple collaborative care in relapse prevention (Simon et al., 2002); given that the study focused on a different population that was not covered by UK studies or other studies ranking higher on the hierarchy of inclusion criteria, this study was also included in the review.

Simple collaborative care

Bosanquet and colleagues (2017) performed a cost-utility analysis alongside a RCT (Bosanquet2017; N=485; at 18 months n=344; cost data available for n=447) that compared simple collaborative care in addition to usual primary care versus primary care alone for older adults who screened positive for major depression in the UK. The perspective of the analysis was the NHS and PSS. Healthcare costs consisted exclusively of intervention and primary care costs. National unit costs were used. The outcome measure was the QALY estimated based on SF-6D ratings (UK tariff). The duration of the analysis was 18 months.

Simple collaborative care was found to be more effective and more costly than usual (primary) care alone, with an ICER of £26,535/QALY (uplifted to 2015 prices). The probability of simple collaborative care being cost-effective at the NICE lower (£20,000/QALY) and upper (£30,000/QALY) cost effectiveness threshold was 0.39 and 0.55, respectively. When only participants who engaged with 5 or more sessions of collaborative care were included in the analysis, the ICER fell at £10,075/QALY. The study is directly applicable to the UK context but is characterised by potentially serious limitations, mainly the inclusion of intervention and primary care costs only.

Green and colleagues (2014) conducted a cost-utility analysis alongside a RCT (Richards2013; N=581, efficacy data available for n=466; resource use data available for n=447) that compared simple collaborative care in addition to usual primary care versus primary care alone for adults with depression in the UK. The perspective of the analysis was the NHS and personal social services (PSS); a broader perspective that included informal care costs and service user expenses was considered in a sensitivity analysis. Healthcare costs consisted of intervention costs, staff time (such as GP, mental health nurse, mental health worker, psychiatrist, psychologist), other outpatient and inpatient care, day care, walk-in-centre, and A&E. National unit costs were used. The outcome measure was the QALY estimated based on EQ-5D ratings (UK tariff); QALY estimates based on the SF-6D (UK tariff) were used in sensitivity analysis. The duration of the analysis was 12 months.

Simple collaborative care was found to be more effective and more costly than usual (primary) care alone, with an Incremental Cost Effectiveness Ratio (ICER) of £15,092/QALY (uplifted to 2015 prices). The probability of simple collaborative care being cost-effective at the NICE lower (£20,000/QALY) and upper (£30,000/QALY) cost effectiveness threshold was 0.58 and 0.65, respectively. Results were robust to multiple imputation of missing data, use of SF-6D utility values, and use of alternative collaborative care costs. The study is directly applicable to the UK context and is characterised by minor limitations.

Lewis and colleagues (2017) also conducted a cost-utility analysis alongside a RCT (Lewis2017; N=705, complete data for economic analysis n=448) that compared simple collaborative care in addition to usual primary care versus primary care alone for older adults who screened positive for subthreshold depression in the UK. The perspective of the analysis was the NHS and PSS. Healthcare costs consisted exclusively of intervention and primary care costs. National unit costs were used. The outcome measure was the QALY estimated based on EQ-5D ratings (UK tariff). The duration of the analysis was 12 months.

Update 2017

Simple collaborative care was found to be more effective and more costly than usual (primary) care alone, with an ICER of £9,827/QALY (uplifted to 2015 prices). The probability of simple collaborative care being cost-effective at the NICE lower (£20,000/QALY) and upper (£30,000/QALY) cost effectiveness threshold was 0.92 and 0.97, respectively. Accounting for the true observed case manager contact rate (rather than the expected contact rate that was used in the base-case analysis), the ICER fell at £3,395/QALY. The study is directly applicable to the UK context but is characterised by potentially serious limitations, mainly the high attrition that was markedly greater in the collaborative care arm, and the consideration of intervention and primary care costs only.

Simon and colleagues (2002) assessed the cost effectiveness of simple collaborative care versus usual care alongside a RCT (Katon2001; N=386, 82% completed all follow-up assessments and 98% remained enrolled throughout the follow-up period) that compared simple collaborative care with treatment as usual for adults with a history of either recurrent major depression or dysthymia that had recovered from a depressive episode following antidepressant treatment in primary care in the US. The study, which adopted a 3rd party payer perspective, considered costs of medication, staff time, as well as costs of any inpatient and outpatient services for mental health or general medical care; local prices were used. The outcome measure was the number of depression-free days, defined as days with a Hopkins Symptoms Checklist (HSCL) depression score ≤ 0.5; days with a HSCL score above 0.5 but < 2 were considered as being 50% depression free. The time horizon of the analysis was 12 months.

Simple collaborative care was found to be more effective and more costly than usual care, with an ICER of $1 per depression-free day (95%CI -$134 to $344, 1998 US$), which translates to £1.1 per depression free day in 2015 prices. The study is only partially applicable to the NICE decision-making context as it was conducted in the US and does not use the QALY as the outcome measure, which requires judgement on whether the additional benefit is worth the extra cost. It is also characterised by potentially serious limitations, resulting mainly from the fact that analyses of clinical data included only those completing all blinded follow-up assessments; cost analyses included only those remaining enrolled throughout the follow-up period. However, participation in follow-up interviews was significantly greater in the intervention group than in usual care, introducing a possibility of bias.

Update 2017

Complex collaborative care

Morriss and colleagues (2016) assessed the cost-utility of complex collaborative care versus usual secondary mental health care in the UK. The economic analysis was carried out alongside a RCT (Morriss2016; N=187; 84% completed at 6 months, 72% at 12 months and 59% at 18 months). Complex collaborating care comprised secondary outpatient specialist depression services offering tailored integrated pharmacological and psychological (CBT, MBCT and compassion focused therapy, as appropriate) treatment within a collaborative care approach for 12-15 months. The analysis adopted a NHS and PSS perspective. Healthcare costs consisted of intervention costs, primary care (GP surgery and home attendances), inpatient and outpatient (psychiatric or other) care, other staff time (practice - district - community psychiatric nurse, psychotherapist), A&E attendances, and medication. National unit costs were used. The outcome measure was the QALY estimated based on EQ-5D ratings (UK tariff). The duration of the analysis was 18 months.

Complex collaborative care was more effective and more costly than usual secondary mental health care, with an ICER of £43,993/QALY (2015 prices). Controlling for baseline differences and cluster effects, the probability of complex collaborative care being cost-effective exceeded 50% at a cost effectiveness threshold of £42,000/QALY, which is well above the NICE cost effectiveness threshold of £30,000/QALY. The study is directly applicable to the UK context and is characterised by minor limitations.

Two Dutch studies assessed the cost effectiveness of complex collaborative care versus treatment as usual in an occupational setting (Goorden et al., 2014) and in primary care (Goorden et al., 2015). Both studies were conducted alongside RCTs (Vlasveld2012 and Huijbregts 2013). Both analyses adopted a healthcare perspective, with productivity losses being reported separately. Healthcare costs consisted of intervention costs (care manager), other staff time (such as GP, mental health care professional, psychologist/psychiatrist, social worker, occupational therapist), self-help groups, day care, psychiatric inpatient care and medication. National unit costs were used. The outcome measure was the QALY estimated based on EQ-5D ratings (Dutch tariff). The time horizon in both analyses was 12 months.

In the occupational setting, complex collaborative care was found to be less effective and less costly than treatment as usual with an ICER of €14,589/QALY (i.e. a saving of €14,589 for every QALY lost) in 2009 prices (£13,233 in 2015 prices), with 75% of the bootstrapped replications suggesting a lower cost and lower efficacy for complex collaborative care compared with treatment as usual. In contrast, in the primary care setting complex collaborative care was found to be more effective and more costly than treatment as usual, with an ICER of €53,717/QALY in 2013 prices (£49,894 in 2015 prices), and a probability of being cost-effective of 0.20 and 0.70 at a cost effectiveness threshold of £18,580 and £74,300/QALY, respectively. These studies are partially applicable to the UK context. The study conducted at the occupational setting (Goorden et al., 2014) is characterised by minor limitations; the study conducted at the primary care setting (Goorden et al., 2015) is characterised by potentially serious limitations, mainly by the fact that, although the RCT included 150 participants, 93 identified by screening and 47 by GP referral, the cost-utility analysis was based only on the 93 participants that were identified by screening

5.2.2.25 Medication management

No UK studies on the cost effectiveness of medication management for adults with depression were identified by the systematic search of the literature. Following the hierarchy of inclusion criteria regarding country settings, one Dutch study (Bosmans et al., 2007) and one Spanish study (Rubio-Valera et al., 2013) were included in the review.

Bosmans and colleagues (2007) evaluated the cost effectiveness of medication management compared with treatment as usual for adults with depression treated in primary care. The study was undertaken alongside a RCT (Brook2005, N=151; economic analysis based on n=88 completers of both 3- and 6-month follow-ups). The study adopted a societal perspective; costs included intervention, staff time (such as GP, psychologist, social worker, psychiatrist, physiotherapist, community mental healthcare, homeopath), laboratory testing, medication and absenteeism from paid labour. National unit prices were used. The outcome measures were the adherence to antidepressant treatment measured using an electronic pill container and depressive symptoms measured using the HSCL. The time horizon of the analysis was 6 months.

Medication management was found to be more costly and more effective than treatment as usual, with an ICER of €14,900 per extra person with improvement in adherence and €2,550 per point improvement in HSCL (2002 prices; translating into figures of £15,314 and £2,621, respectively, in 2015 prices). The probability of medication management being cost-effective was approximately 0.65 at a cost effectiveness threshold of €50,000 (£51,391 in 2015 prices) per extra person with improvement in adherence. Results were robust to different scenarios such as a per protocol analysis, a change in intervention cost, use of different methodology for estimating indirect costs, and imputation of missing data. The study is partially applicable to the UK decision-making context, as it was conducted in the Netherlands and adopted a societal perspective, including absenteeism costs. Moreover, it did not use the QALY as a measure of outcome, so results required further judgements on whether the intervention is cost-effective. The study was characterised by potentially serious limitations, such as its

Update 2017

short time horizon and the limited sub-sample (out of the randomised sample) it was based on.

Rubio-Valera and colleagues (2013) conducted an economic evaluation of medication management versus treatment as usual for adults with depression treated in primary care. The study was undertaken alongside a RCT (Rubio-Valera2012, N=179; 71% completed at 6 months; n=151 received intervention as allocated). The study adopted a healthcare and a societal perspective; costs included intervention, publicly funded healthcare services (GP, nurse, psychologist, psychiatrist, other specialists, social worker, hospital emergency visits, hospital stay, diagnostic tests, medication), privately funded healthcare services (psychiatrist, psychologist, medical specialist, GP), and absenteeism from paid labour. Regional unit prices were used. The study used 3 outcome measures: adherence to antidepressant treatment measured using electronic pharmacy records; remission of depressive symptoms defined as a reduction in the Patient Health Questionnaire 9-item (PHQ-9) of at least 50%; and the QALY based on EQ-5D ratings and the Spanish tariff. The time horizon of the analysis was 6 months.

Under the healthcare perspective, medication management was more expensive than treatment is usual. It was also more effective in terms of adherence to antidepressant treatment and the QALYs gained. The respective ICERs were €962 per extra adherent service user and €3,592/QALY (2009 prices; translating into figures of £863 and £3,224, respectively, in 2015 prices). However, when remission was used as an outcome, medication management was dominated by treatment as usual, as it was more expensive and less effective. The probability of medication management being cost-effective was 0.71 and 0.76 for WTP £5,385/adherent service user and £26,927/QALY, respectively (2015 prices). Using remission as an outcome, the maximum probability of medication management being cost-effective was only 0.46, irrespective of the cost effectiveness threshold used. Results were robust to different scenarios such as a per protocol or complete case analysis, use of different diagnostic criteria for depression, changes in intervention costs or different methodology used for estimating indirect costs. The study is partially applicable to the UK decision-making context, as it was conducted in Spain. The findings of the study are inconsistent across the outcome measures used (i.e. the study appears to be cost-effective using the QALY, but cost-ineffective using remission as measure of outcome). The study was characterised by potentially serious limitations, mainly its contradictory results, its short time horizon and the use of regional unit costs.

Update 2017

5.2.2.34 Care co-ordination

No studies assessing the cost effectiveness of care co-ordination for adults with depression were identified by the systematic search of the literature.

5.2.2.47 Stepped care

The systematic search of the literature identified one UK study assessing the cost effectiveness of stepped care (Mukuria et al., 2013); another German economic study of stepped care was also included in the economic review of stepped care following the hierarchy of inclusion criteria regarding country settings (Ricken et al., 2011).

Mukuria and colleagues (2013) assessed the cost-utility of stepped care for people with depression or anxiety in the UK, as reflected in the Improving Access to Psychological Therapies (IAPT) service, in addition to treatment as usual, versus treatment as usual alone; the latter comprised GP care, primary care counselling and referral to secondary mental health services. The study was conducted alongside a prospective cohort study with matched sites (N=403), and more than 95% of the study sample included people with a primary diagnosis of depression. The analysis adopted a NHS and social services perspective; productivity losses were assessed separately. Healthcare costs consisted of intervention (staff time, training, equipment, facilities and overheads), other mental

healthcare (psychiatrist, psychologist, community psychiatric nurse, etc.), primary and secondary care, and social care; medication costs were not considered. Unit costs were based on IAPT data and national sources. The outcome measures of the analysis were the proportion of people with a reliable and clinically significant (RCS) improvement on the PHQ-9 and the QALY based on SF-6D ratings (UK tariff); QALYs estimated based on predicted EQ-5D ratings (UK tariff), estimated from SF-6D using an empirical mapping function, were used in sensitivity analysis. The duration of the analysis was 8 months.

IAPT added to treatment as usual was more costly and more effective than treatment as usual alone, with ICERs of £10,363 per additional participant with RCS improvement, £32,384/QALY using the SF-6D and £18,504/QALY using predicted EQ-5D scores (figures uplifted to 2015 prices). The probability of IAPT being cost-effective using SF-6D QALYs was less than 0.40 at a cost effectiveness threshold of £30,000/QALY; using QALYs estimated based on predicted EQ-5D ratings the probability of IAPT being cost-effective was 0.38 and 0.53 at cost effectiveness thresholds of £20,000 and £30,000/QALY, respectively. Using national unit costs instead of IAPT financial data resulted in an ICER of £4,171 per additional participant achieving RCS improvement and £13,036/QALY using SF-6D ratings. It is noted that NICE recommends use of EQ-5D for the estimation of QALYs in adults.

The study is directly applicable to the UK context and is characterised by potentially serious limitations such as its short time horizon, its study design, the sensitivity of results to unit costs of IAPT, the low response rate at recruitment (403 out of 3,391, 11.9%); and the fact that the IAPT service was assessed over the first 2 years of establishment, therefore costs associated with learning effects were likely.

Ricken and colleagues (2011) assessed the cost effectiveness of stepped care in an inpatient setting, comprising a standardised stepwise drug treatment regimen, compared with inpatient treatment as usual, for adults with depression in Germany, by conducting an economic analysis alongside a RCT (Bauer2009, N=148; completers n=103). The analysis adopted a 3rd party payer perspective and included only medication and hospitalisation costs, priced using national unit costs. The measure of outcome was remission, defined as a Bech–Rafaelsen-Melancholia-Scale (BRMS) score <7. The duration of the analysis was the time from enrolment to study endpoint, i.e. dropout or remission.

Stepped care was found to dominate treatment as usual, as it was more effective and less costly. The study is partially applicable to the UK as it was conducted in Germany. The study has not used the QALY, but results were straightforward to interpret as the intervention was dominant. The study is characterised by potentially serious limitations, such as the consideration of hospitalisation and medication costs only, and the duration of the analysis, from enrolment to study endpoint, which did not allow estimation of re-hospitalisation costs, costs incurred after hospital discharge, etc.

Update 2017

5.2.2.58 Integrated care pathways

No UK studies assessing the cost effectiveness of integrated care pathways were identified by the systematic literature search. Following the hierarchy of inclusion criteria regarding study settings, two US economic studies in this area were included in the review (Pyne et al., 2015; Wiley-Exley et al., 2009).

Pyne and colleagues (2015) assessed the cost effectiveness of integrated local primary care (primary care liaison) co-ordinated by on-site nurse depression care managers versus off-site specialists, for adults with depression in the US. The analysis was undertaken alongside a RCT (Dobscha2006, N=364; 87% completed at 6 months, 79% at 12 months and 78% at 18 months). The analysis adopted a healthcare and service users' perspective and included intervention costs, inpatient and outpatient care, emergency room care, medication, and also service users' time and mileage. The study utilised regional sources for unit costs, with national unit costs being used in a secondary analysis. The measures of outcome were the number of depression-free days derived from HSCL (score ≤ 0.5 indicated depression-free

day, ≥ 1.7 full symptoms and intermediate severity scores were assigned a value between depression-free and fully symptomatic by linear interpolation); and the QALY, estimated based on the SF-12/SF-6D algorithm (UK tariff). The duration of the analysis was 18 months.

Integrated care by off-site managers care was more effective and more costly than integrated care managed by on-site managers, with an ICER of $36,033/QALY using regional costs or $28,126/QALY using national costs (2009 prices; translated into £25,875 and £20,197/QALY, respectively, in 2015 prices). The probability of off-site integrated care being cost-effective was 0.86 at a cost effectiveness threshold of $50,000/QALY (£35,901/QALY in 2015 prices). Results per depression-free day did not include inpatient care costs and therefore these are not reported here. The study is partially applicable to the UK as it was conducted in the US, and is characterised by minor limitations.

Wiley-Exley and colleagues (2009) evaluated the cost effectiveness of integrated care compared with primary care with a referral system to specialist care for older adults with depression in the US. The study, which was conducted alongside a RCT (N=840), analysed 4 different combinations of populations and settings: people major and minor depression (full sample) in the Veteran Affairs (VA) setting (n=365), full sample outside VA (n=475); people with major depression within VA (n=214), and people with major depression outside VA (n=302). The analysis adopted a healthcare and service users' and carers' perspective and included intervention costs, outpatient and inpatient care, nursing home, rehabilitation, emergency room, medication, service users' and caregivers' time and travel costs. National unit costs were used. The study included various measures of outcome, such as the CES-D score; the number of depression-free days derived from CES-D; the number of QALYs estimated based on depression-free days, using utility weights of health=1, depression=0.59; the number of QALYs estimated based on SF-36, using preferences for matched vignettes created following cluster analysis of SF-12 mental and physical component scores, elicited by US service users with depression using SG. Only results for the latter are reported here (full results of the study are provided in the study's evidence table in Appendix Q). The time horizon of the analysis was 6 months.

Integrated care was found to dominate usual primary care in the full sample (major and minor depression), VA setting. It was more costly and more effective than usual primary care regarding the full sample outside VA setting and major depression sample in the VA setting, with ICERs of £84,566/QALY and £52,395/QALY, respectively (2015 prices). It was less effective and less costly than usual primary care in the major depression sample, outside the VA setting, with an ICER of £70,902/QALY (saving per QALY lost).

The probability of integrated care being cost-effective was more than 0.70 for any cost effectiveness threshold only in the full sample and VA setting. The probability of integrated care being cost-effective was low at levels of willingness to pay that corresponded to NICE cost effectiveness thresholds. The study is partially applicable to the UK as it was conducted in the US, and is characterised by potentially serious limitations, including the short time horizon and the contradictory results across sub-analyses.

Update 2017

5.2.3 Clinical evidence statements

5.2.3.1 Collaborative care

- Very low quality evidence from up to 47 RCTs (k=3-47, n=up to 4539) showed that both simple and complex collaborative care models have a small beneficial effect on depression symptoms at 6 months, and that at 12 months collaborative care overall and simple collaborative care specifically have a small beneficial effect, whilst complex collaborative care had a stronger, clinically important, but not statistically significant beneficial effect over control.
- Very low-low quality evidence from 3 different RCTs (k=3, n=214-1041) showed no difference in remission rates at 6 month follow-up between those provided with simple

collaborative care or control, but a clear benefit of both simple and complex collaborative care at 12 month follow-up.

- Very low-low quality evidence from 10 RCTs (k=10, n=3278) showed a clear benefit of collaborative care overall, and of both simple and complex within that, on response rates at 12 month follow-up when compare with control.
- Very low quality evidence from up to 31 RCTs (k=4-31, n=up to 3618) showed greater antidepressant use in the collaborative care condition than control at 6 month follow-up, with this effect being slightly more pronounced in the simple collaborative care than complex collaborative care conditions, and also at 12 month follow-up in collaborative care overall and complex collaborative care. At 12 month follow-up however, there was a clinically important but not statistically significant increase in antidepressant use in simple collaborative care compared with control.
- Low quality evidence from 1 RCT (k=1, n=132) showed a clinically important but not statistically significant increase in remission rates at 12 month follow-up in patients provided with patient-centred compared with standard simple collaborative care.
- Low-moderate quality evidence from 1 RCT (k=1, n=287-318) showed greater response rates at both 6 and 12 month follow-up in patients treated with practice-based collaborative care compared with tele-based collaborative care.

5.2.3.2 Stepped care

- Very low-low quality evidence from 3 different RCTs (k=3, n=148-201) showed a clinically important but not statistically significant benefit of stepped care over control on remission, depressive symptoms as measured on the PHQ-9 and antidepressant use at 6 months.

5.2.3.3 Care co-ordination

- Very low-low quality evidence from up to 4 RCTs (k=3-4) showed no benefit of care co-ordination over control on mean change in depression scores at 6 months, however there was a clinically important but not statistically significant increase in antidepressant adherence in the care coordination condition.

Update 2017

5.2.3.4 Medication management

- Very low quality evidence from up to 9 RCTs (k=1-9, n=219 and over) showed no benefit of medication management over control on mean change in depression scores at 6 month follow-up and a clinically important but not statistically significant benefit of medication management at 12 month follow-up.
- Very low quality evidence from 4 RCTs showed a clinically important but not statistically significant benefit of medication management over control on antidepressant use at 6 month follow-up.

5.2.3.5 Integrated care pathways

- Very low-low quality evidence from 3 different RCTs (k=1-2) showed no difference overall, or for integrated care versus speciality referral system specifically, in mean change in depression scores at 6 months, however there was a clinically important but not statistically significant decrease in depression scores in patients in the integrated care condition compared with control at 6 months which was not maintained at 12 months. Furthermore there was a clinically important but not statistically significant increase in antidepressant adherence in integrated care compared with a control condition.

5.2.3.6 Service delivery models for relapse prevention

- Very low-low quality evidence from 2 different RCTs (k=1-1, n=136-386) showed no benefit of simple collaborative care over control on depressive symptoms at 6 month

follow-up, or on relapse rates at 12 month follow-up, and a clinically important but not statistically significant benefit of control over stepped care on relapse prevention at 12 month follow-up.

5.2.44 Economic evidence statements

5.2.4.15 Collaborative care

- Evidence from 3 UK economic evaluations conducted alongside RCTs (N = 1,771; complete data for economic analysis n=1341) suggest that simple collaborative care is possibly a cost-effective model for delivering services to adults with depression. This evidence is directly applicable to the UK context and is coming from one study with minor and two studies with potentially serious methodological limitations.
- Evidence from 1 US study conducted alongside a RCT (N=386) suggests that simple collaborative care aiming at relapse prevention may be cost-effective in adults with depression that is in remission. This evidence is partially applicable to the NICE decision-making context as it comes from a US study and is not using the QALY as the outcome measure. The study is characterised by potentially serious methodological limitations.
- Evidence from 1 UK study conducted alongside a RCT (N=187) suggests that complex collaborative care is not cost-effective compared with usual secondary mental health care for adults with depression. This evidence is directly applicable to the UK context and is characterised by minor limitations.
- Evidence from 2 Dutch studies conducted alongside RCTs (N=219) suggest that complex collaborative care is unlikely to be cost-effective compared with treatment as usual in adults with depression. This evidence is partially applicable to the NICE decision-making context as the studies were conducted in the Netherlands and utility values were based on EQ-5D ratings using the Dutch tariff. One study is characterised by minor limitations and the other study by potentially serious limitations.

Update 2017

5.2.4.26 Medication management

- Evidence from 1 Dutch and 1 Spanish study conducted alongside RCTs (N=330) is inconclusive regarding the cost effectiveness of medication management for adults with depression. This evidence is partially applicable to the NICE decision-making context as the studies were conducted outside the UK. The Dutch study adopted a societal perspective and did not use the QALY as the measure of outcome, therefore further judgements were required in order to assess the cost effectiveness of medication management. The Spanish study included the QALY as one of the measures of outcome, based on EQ-5D ratings and the Spanish values. Both studies are characterised by potentially serious limitations.

5.2.4.36 Care co-ordination

- No evidence on the cost effectiveness of care co-ordination for adults with depression is available.

5.2.4.49 Stepped care

- Evidence from 1 UK study conducted alongside a cohort study with matched sites (N=403) and 1 German study conducted alongside a RCT (N=148) suggests that stepped care might be cost-effective for adults with depression. This evidence is directly applicable (UK study) and partially applicable (German study) to the NICE decision-making context. Both studies are characterised by potentially serious limitations.

5.2.4.5 Integrated care pathways

- Evidence from 1 US study conducted alongside a pragmatic RCT (N=364) suggests that integrated care managed by off-site managers may be more cost-effective than on-site managed integrated care. The evidence is partially applicable to the NICE decision-making context (US study, QALYs based on SF-12/SF-6D algorithm - UK tariff) and is characterised by minor limitations.
- Evidence from 1 US study conducted alongside a multi-site pragmatic RCT (N=840) is inconclusive regarding the cost effectiveness of integrated care compared with usual primary care that includes a referral system to specialist care. The evidence is partially applicable to the NICE decision making context (US study, QALYs based on SF-36, using preferences for matched vignettes created following cluster analysis of SF-12 mental and physical component scores, elicited by US service users with depression using SG) and is characterised by minor limitations.

5.2.5 From evidence to recommendations

5.2.5.1 Relative values of different outcomes

The GC identified depression symptomology (6 months) and response, remission and relapse (12 months) to be the critical outcomes for this question. Service utilisation and resource use were identified as important outcomes.

Evidence was available for all outcomes of interest for the collaborative care dataset, and for relapse prevention from the stepped care and collaborative care datasets. A number of different care models did not have available data on the outcomes of remission and response. Therefore when considering the evidence the GC placed the greatest emphasis on depression symptoms and resource use (antidepressant use), as these provided the best point of comparison across different interventions.

Update 2017

5.2.5.2 Trade-off between clinical benefits and harms

In developing the recommendations for service organisation the GC were mindful of the problems that people with depression and, in particular, people with more severe depression have in accessing and engaging with services in both primary and secondary care. The GC therefore considered the evidence on collaborative care and decided that the provision of a simple model of collaborative care could be effective in ensuring both greater engagement with and uptake of services for people with more severe depression. Also, given that engagement issues are even greater in older adults, in particular those with physical health problems, and that there was evidence of the cost-effectiveness of collaborative care in older people with chronic physical health problems the GC agreed to recommend collaborative care for this group of people.

The GC were aware of the importance of medication adherence, in particular, for people with severe and chronic depression and did consider the evidence on medication management. They noted the very limited evidence for medication management and that for most people the delivery of care in a collaborative, multidisciplinary manner was more effective at promoting medication adherence. Therefore the GC agreed to recommend that medication management should not be provided as a separate intervention.

The GC acknowledged that for more severe depression with multiple complicating problems or significant coexisting conditions there was no direct evidence to guide the development of recommendations. The GC were however aware of the very significant burden people with severe and complex depression face and the burden this represents for families and carers. Such high levels of need are best met by specialist services within specialist secondary care services. The GC therefore drew on their expert knowledge and experience of specialist services and used informal consensus to develop a series of recommendations on who might

benefit for specialist services; how these services should be co-ordinated and what the nature of the co-ordination of the services should involve. In the view of the GC referral to specialist services would ensure that this population receives appropriate care for their condition, leading to improved outcomes and likely cost-savings from reduction in the need for costly care further down the care pathway in the absence of a clear referral process. The GC were of the view that the development of a comprehensive multidisciplinary care plan will allow more timely, appropriate and potentially cost-effective planning and delivery of care to people with more severe depression with multiple complicating problems or significant coexisting conditions, that is targeted to their specific needs and thus can result in cost-savings that offset, fully or partially, the costs associated with development of the care plan. In contrast, lack of a detailed care plan may lead to sub-optimal, less clinically and cost-effective care pathways and inappropriate treatments, ultimately leading to sub-optimal outcomes for the person and higher healthcare costs.

The GC considered that effective service delivery models would enhance clinical outcomes by improved engagement with effective interventions and thereby improve outcomes in terms of depressive symptomology and response, remission and relapse. They noted that there was evidence from a number of UK and international trials that there were clinical benefits associated with the use of collaborative care. There was more limited clinical evidence to support the use of a stepped care model for the provision of care. The evidence for medication management, integrated care pathways and care co-ordination was very limited. The GC took the view that the potential harms would be poorer engagement with services, poorer adherence whilst in treatment and consequently poorer outcomes. These models of care could interfere with establish care pathways with which service users are familiar and therefore could result in poorer access, uptakes and outcomes.

5.2.5.3 Trade-off between net health benefits and resource use

Collaborative care

There is evidence from 3 UK economic evaluations conducted alongside RCTs that simple collaborative care is potentially a cost-effective model for delivering services to adults with depression; some of this evidence is characterised by potentially serious methodological limitations. Another UK study indicated that complex collaborative care is unlikely to be cost-effective for this population. These conclusions about simple and complex collaborative care are supported by non-UK evidence, derived from studies conducted in the US and the Netherlands.

The GC noted that, overall, the published economic evidence indicated that simple collaborative care is likely to be a cost-effective model for delivering services to adults with depression; in contrast, more resource-intensive complex collaborative care is unlikely to be cost-effective compared with usual care.

Medication management

The GC noted that no UK evidence was available and non-UK evidence did not provide any substantial support for the cost effectiveness of medication management as an independent care model for adults with depression.

Stepped care

Evidence from one UK study and one German study suggested that stepped care might be cost-effective for adults with depression. Both studies were characterized by potentially serious limitations. The GC noted, based on the evidence, that stepped care might be cost effective for adults with depression.

Update 2017

Integrated care pathways

Two US studies on the cost effectiveness of integrated care pathways were identified. Both studies were assessed as having minor limitations. The GC noted that the published evidence was inconclusive about the cost effectiveness of integrated care

Care co-ordination

No evidence was identified on the cost-effectiveness of care co-ordination.

The GC acknowledged that referring people with more severe depression and multiple complicating problems (such as unemployment, poor housing or financial problems) or significant coexisting conditions to specialist mental health services is likely to incur additional costs compared with no referral. However they agreed that the number of people affected would be small and any additional costs were likely to be offset by cost-savings resulting from more appropriate care for this population following referral (compared with treatment in primary care settings), leading to improved outcomes and reduction in the need for potentially costly care further down the care pathway.

5.2.5.45 Quality of evidence

Collaborative care

Very low quality evidence from up to 47 RCTs (k=3-47, n=up to 4539) showed that both simple and complex collaborative care models have a beneficial effect on depression symptoms. Similar very low to low quality evidence showed no difference in remission rates at 6 month follow-up between those provided with simple collaborative care or control, but a clear benefit of both simple and complex collaborative care at 12 month follow-up. Very low to low quality evidence from 10 RCTs (k=10, n=3278) showed a clear benefit of collaborative care overall, and of both simple and complex within that, on response rates at 12 month follow-up when compare with control. Very low to low quality evidence showed no benefit of simple collaborative care on depressive symptoms at 6 month follow-up, or on relapse rates at 12 month follow-up.

Stepped care

For stepped care, very low to low quality evidence showed a clinically important but not statistically significant benefit of stepped care over control on remission, depressive symptoms and antidepressant use at 6 months.

Medication management

For medication management, very low quality evidence showed no benefit of medication management over control on change in depression scores at 6 month follow-up and a clinically important but not statistically significant benefit of medication management at 12 month follow-up. Very low quality evidence showed a clinically important but not statistically significant benefit of medication management over control on antidepressant use at 6 month follow-up. No evidence was identified of any harms for this intervention.

Integrated care pathways and care co-ordination

Very low to low quality evidence showed no benefit of care co-ordination over control on depression scores at 6 months. For integrated care pathways there was very low to low quality evidence showing no difference in depression scores at 6 months. There was a clinically important but not statistically significant decrease in depression scores in patients in the integrated care pathway compared with control at 6 months.

Update 2017

5.3 Recommendations

Collaborative care

6. **Consider collaborative care for all older people with depression, in particular if they have significant physical health problems or social problems. [new 2017]**

7. **Consider collaborative care as a method for the delivery of care for people with more severe depression. [new 2017]**

8. **Ensure that collaborative care for people with more severe depression covers:**
   - patient-centred assessment and engagement
   - symptom measurement and monitoring
   - medication management
   - active follow-up by a designated case manager
   - delivery of psychological and psychosocial interventions within a structured protocol, for example stepped care
   - taking any relevant physical health problems into account
   - regular liaison with primary and secondary care colleagues
   - supervision of practitioner(s) by an experienced mental health professional. [new 2017]

Specialist care planning

9. **Refer people to specialist mental health services for a programme of coordinated multidisciplinary care if they have:**
   - more severe depression with multiple complicating problems, for example, unemployment, poor housing or financial problems, or
   - significant coexisting conditions. [new 2017]

10. **Ensure multidisciplinary care plans for people with more severe depression with multiple complicating problems, or significant coexisting conditions:**
    - are developed together with the person, their GP and other relevant people involved in their care (with the person's agreement)
    - set out the roles and responsibilities of all health and social care professionals involved in delivering the care
    - include information about 24-hour support services, and how to contact them
    - include a crisis plan that identifies potential crisis triggers, and strategies to manage those triggers
    - are updated if there are any significant changes in the person's needs or condition
    - are reviewed at agreed regular intervals
    - include medication management (a plan for starting, reviewing and discontinuing medication). [new 2017]

Update 2017

5.4 Review question

- For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 23. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Update 2017

Table 23: Clinical review protocol summary for the review of settings for care of adults with depression

Component	Description
Review question	For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care? (RQ 1.2)
Population	Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by depression scale score for subthreshold and other groups
Intervention(s)	Settings for the delivery of care, which may include: • Primary care • Crisis resolution and home treatment teams • Inpatient setting • Acute psychiatric day hospital care • Non-acute day hospital care and recovery centres • Specialist tertiary affective disorders settings • Community Mental Health Teams • Residential services
Comparison	• Any other setting for the delivery of care
Critical outcomes	• Depression symptomology (e.g. mean endpoint score or change in depression score from baseline) • Response (e.g. reduction of at least 50% from the baseline score on depression scale) • Remission (e.g. score below a certain a threshold on a depression scale) • Relapse (number of people who relapsed)
Important but not critical outcomes	• Service utilisation/resource use (e.g. antidepressant use)
Study design	• Systematic reviews of RCTs • RCTs • Cluster RCTs

5.4.1 Clinical evidence

The higher order question addressed by this review question is as follows:

- Is there anything about the general management of care that should be done differently when delivered in different settings?

Trials of interventions delivered in certain settings will recruit populations relevant to that setting. However, ideally in order to address this question we would want trials that randomise the same population to different settings for the delivery of care. Evidence for this is limited and the review approach differed slightly depending on the best evidence available, the approach and evidence will be presented below for each setting as follows: primary care; crisis resolution and home treatment teams; inpatient care; acute psychiatric day hospital

care; non-acute day hospital care and recovery centres; specialist tertiary affective disorders settings; community mental health teams (CMHTs); residential services.

5.4.1.13 Primary care

No RCT evidence was identified that specifically addressed this setting. Therefore the GC considered indirect evidence in the form of sub-analyses of the NMA dataset (acute treatment of depressive episodes).

61 RCTs were included in this analysis, 21 in primary care settings and 40 in secondary care settings. Seven comparisons addressing different treatment options were possible with this data; these were i) amitriptyline versus placebo, ii) IPT versus TAU/waitlist, iii) counselling versus TAU/waitlist, iv) behavioural therapies versus TAU/waitlist, v) cognitive and cognitive-behavioural therapies versus TAU/waitlist, vi) self-help versus TAU/waitlist and vii) self-help with support versus TAU/waitlist. See Table 24, Table 25 and Table 26 for study characteristics and Appendix M for forest plots.

Primary versus secondary care differences were examined for outcomes that had more than one study in each subgroup. No significant subgroup differences for primary care compared to secondary care were found for the amitriptyline versus placebo comparison (Discontinuation for any reason: $Chi^2 = 0.05$, df = 1, p = 0.82; Discontinuation due to side effects: $Chi^2 = 0.06$, df = 1, p = 0.80; Depression symptomatology had <2 studies per subgroup). All outcomes had less than two studies per subgroup for the IPT versus treatment as usual or waitlist and the counselling versus treatment as usual comparisons. No significant subgroup differences for primary care compared to secondary care were found for the behavioural therapies versus treatment as usual or waitlist comparison (Depression symptomatology: $Chi^2 = 0.26$, df = 1, p = 0.61; Discontinuation for any reason: $Chi^2 = 0.02$, df = 1, p = 0.89; Remission had <2 studies per subgroup). No significant subgroup differences for primary care compared to secondary care were found for the cognitive and cognitive behavioural therapies versus treatment as usual or waitlist comparison (Depression symptomatology: $Chi^2 = 0.56$, df = 1, p = 0.46; Remission: $Chi^2 = 0.32$, df = 1, p = 0.57; Discontinuation for any reason: $Chi^2 = 0.05$, df = 1, p = 0.82). There was evidence for a statistically significant difference between primary and secondary care subgroups for self-help (without support) versus treatment as usual or waitlist on depression symptomatology ($Chi^2 = 4.94$, df = 1, p = 0.03), with evidence for statistically significant benefits of self-help in both primary care and secondary care studies but larger benefits observed in secondary care (SMD -0.59 [-0.84, -0.34]) than in primary care (SMD -0.27 [-0.40, -0.13]). However, although the overall effect size was larger in the secondary care subgroup, there were also more included studies and participants (K=13 and N=1479 compared to K=3 and N=832) and heterogeneity was considerably higher (I^2=79% compared to I^2=0%), so no clear conclusions are possible based on this finding. No significant subgroup differences were found for the only other permissible outcome for this self-help (without support) versus treatment as usual or waitlist comparison (Discontinuation for any reason: $Chi^2 = 0.02$, df = 1, p = 0.89; Remission had <2 studies per subgroup). For the self-help with support versus treatment as usual or waitlist comparison, there were no statistically significant subgroup differences for efficacy outcomes (Depression symptomatology: $Chi^2 = 2.38$, df = 1, p = 0.12; Remission: $Chi^2 = 2.06$, df = 1, p = 0.15). However, there was a statistically significant difference between primary care and secondary care subgroups on the discontinuation for any reason outcome in the self-help with support versus treatment as usual or waitlist comparison ($Chi^2 = 5.79$, df = 1, p = 0.02). Visual inspection of the forest plot reveals a neither clinically important nor statistically significant effect of self-help with support relative to treatment as usual or waitlist on discontinuation in primary care studies (K=5; N=1409). However, in secondary care studies (K=5; N=382) drop-out is significantly greater (over twice as high) in the self-help with support arm relative to treatment as usual or waitlist, suggesting there may be more issues with the acceptability of self-help with support in secondary care compared to in primary care.

Update 2017

Table 24: Study information table for trials included in the sub-analysis of primary care versus secondary care (part 1 – pharmacological interventions)

	Amitriptyline versus placebo
Total no. of studies (N randomised)	*Primary care* 2 (150) *Secondary care* 11 (1217)
Study ID	*Primary care* Mynors-Wallis 1995[1] Thomson 1982[2] *Secondary care* Amsterdam 1986[3] Bakish 1992a[4] Bakish 1992b[5] Gelenberg 1990a[6] Hicks 1988[7] Hollyman 1988[8] Lydiard 1997[9] McCallum 1975[10] Rickels 1985[11] Spring 1992[12] Wilcox 1994[13]
Country	*Primary care* UK[1,2] *Secondary care* US[3,6,7,9,11,12,13] Canada[4,5] UK[8] Australia[10]
Baseline depression severity	*Primary care* Less severe[1,2] *Secondary care* Less severe[3,4,5,8,9,10] More severe[6,7,11,12,13]
Age (mean)	*Primary care* 37.1[1] Median age=33 years[2] *Secondary care* 41[3] 43.4[4] 43.0[5] NR[6,8] 41.5[7,10] 39.6[9] 39[11] 34.9[12] 40[13]
Sex (% female)	*Primary care* 74[1] NR[2] *Secondary care* 34[3]

	Amitriptyline versus placebo
	60[4] 43[5] 69[6] NR[7] 83[8,10] 68[9] 86[11] 68[12] 70[13]
Ethnicity (% BME)	*Primary care* NR[1,2] *Secondary care* NR[3,4,5,6,7,8,10,11,12,13] 5[9]
Intervention	*Primary care* Amitriptyline 50-150mg/day[1] Amitriptyline 75-150mg/day[2] *Secondary care* Amitriptyline 100-300mg/day[3] Amitriptyline 50-150mg/day[4,5,6,9] Amitriptyline 25-300m/day[7] Amitriptyline 75-175mg/day[8] Amitriptyline 150mg/day[10] Amitriptyline 50-225mg/day[11] Amitriptyline 50-350mg/day[12] Amitriptyline 60-300mg/day[13]
Comparison	*Primary care* Pill placebo *Secondary care* Pill placebo
Notes: [1]Mynors-Wallis 1995; [2]Thomson 1982; [3]Amsterdam 1986; [4]Bakish 1992a; [5]Bakish 1992b; [6]Gelenberg 1990a; [7]Hicks 1988; [8]Hollyman 1988; [9]Lydiard 1997; [10]McCallum 1975; [11]Rickels 1985; [12]Spring 1992; [13]Wilcox 1994 Mynors-Wallis 1995 and Lydiard 1997 are three-armed trials but where possible the demographics reported here are for only the two relevant arms.	

Table 25: Study information table for trials included in the sub-analysis of primary care versus secondary care (part 2 – formal psychological interventions)

	IPT versus TAU/waitlist	Counselling versus TAU	Behavioural therapies versus TAU/waitlist	Cognitive and cognitive-behavioural therapies versus TAU/waitlist
Total no. of studies (N randomised)	*Primary care* 2 (265) *Secondary care* 3 (314)	*Primary care* 2 (194) *Secondary care* 1 (453)	*Primary care* 2 (202) *Secondary care* 3 (259)	*Primary care* 7 (639) *Secondary care* 7 (566)
Study ID	*Primary care* Beeber 2010[1] Schulberg 1996[2]	*Primary care* Scott 1992[6] Ward 2000[7]	*Primary care* Dalgard 2006[9] Ekers 2011[10]	*Primary care* Cramer 2011[14]

	IPT versus TAU/waitlist	Counselling versus TAU	Behavioural therapies versus TAU/waitlist	Cognitive and cognitive-behavioural therapies versus TAU/waitlist
	Secondary care Lemmens 2015 /2016[3] Swartz 2008[4] Van Schaik 2006[5]	*Secondary care* MacPherson 2013[8]	*Secondary care* Gawrysiak 2009[11] McIndoo 2016[12] Spek 2007[13]	Dwight-Johnson 2011[15] Laidlaw 2008[16] Miranda 2003[17] Scott 1992[18] Serfaty 2009[19] Ward 2000[20] *Secondary care* Kohtala 2015[21] Lemmens 2015 /2016[3] Losada 2015[22] Mohr 2011[23] Naeem 2015[24] Scott 1997[25] Selmi 1990[26]
Country	*Primary care* US[1,2] *Secondary care* Netherlands[3,5] US[4]	*Primary care* UK[6,7] *Secondary care* UK[8]	*Primary care* Norway[9] UK[10] *Secondary care* US[11,12] Netherlands[13]	*Primary care* UK[14,16,18,19,20] US[15,17] *Secondary care* Finland[21] Netherlands[3] Spain[22] US[23,26] Pakistan[24] UK[25]
Baseline depression severity	*Primary care* Less severe[1,2] *Secondary care* More severe[3] Less severe[4,5]	*Primary care* Less severe[6] More severe[7] *Secondary care* Less severe[8]	*Primary care* Less severe[9] More severe[10] *Secondary care* Less severe[11,12,13]	*Primary care* Less severe[14,15,16,17,18] More severe[19,20] *Secondary care* Less severe[21,22,23,24,25,26] More severe[3]
Age (mean)	*Primary care* 26.4[1] 37.9[2] *Secondary care* 40.0[3] 42.8[4] 67.9[5]	*Primary care* 33.9[6] 38[7] *Secondary care* 43.5[8]	*Primary care* 47.3[9] 44.7[10] *Secondary care* 18.4[11] 19.2[12] 54.5[13]	*Primary care* 42.5[14] 39.8[15] 74.1[16] 29.7[17] 30.2[18] 73.6[19] 36.5[20] *Secondary care* 46.2[21] 40.0[3] 61.8[22] 55.9[23] 31.7[24]

	IPT versus TAU/waitlist	Counselling versus TAU	Behavioural therapies versus TAU/waitlist	Cognitive and cognitive-behavioural therapies versus TAU/waitlist
				41.0[25] 27.8[26]
Sex (% female)	*Primary care* 100[1] 85[2] *Secondary care* 72[3] 100[4] 69[5]	*Primary care* 78[6] 77[7] *Secondary care* 75[8]	*Primary care* 76[9] 62[10] *Secondary care* 80[11] 63[12] 61[13]	*Primary care* 100[14,17] 78[15,18] 73[16] 82[19] 76[20] *Secondary care* 79[21] 65[3] 84[22] 9[23] 60[24] 67[25,26]
Ethnicity (% BME)	*Primary care* 100[1] NR[2] *Secondary care* NR[3,4,5]	*Primary care* NR[6] 9[7] *Secondary care* NR[8]	*Primary care* NR[9,10] *Secondary care* 30[11] 27[12] NR[13]	*Primary care* 11[14] NR[15,16,18] 94[17] 7[19] 10[20] *Secondary care* NR[3,21,22,24,25] 21[23] 0[26]
Intervention	*Primary care* Interpersonal psychotherapy (IPT) [1,2] *Secondary care* Interpersonal psychotherapy (IPT) [3,4,5]	*Primary care* Directive counselling[6] Non-directive counselling[7] *Secondary care* Non-directive counselling[8]	*Primary care* Coping with Depression course (group) [9] Behavioural activation (BA) [10] *Secondary care* Behavioural activation (BA) [11,12] Coping with Depression course (group) [13]	*Primary care* CBT group (under 15 sessions)[14] CBT individual (under 15 sessions)[15,17,18,19,20] CBT individual (over 15 sessions)[16] *Secondary care* Third-wave cognitive therapy individual[21] CBT individual (over 15 sessions)[3,23] CBT individual (under 15 sessions) and Third-wave cognitive therapy individual arms combined[22] CBT individual (under 15 sessions)[24,25,26]

	IPT versus TAU/waitlist	Counselling versus TAU	Behavioural therapies versus TAU/waitlist	Cognitive and cognitive-behavioural therapies versus TAU/waitlist
Comparison	*Primary care* Treatment as usual[1,2] *Secondary care* Waitlist[3] Treatment as usual[4,5]	*Primary care* Treatment as usual[6,7] *Secondary care* Treatment as usual[8]	*Primary care* Treatment as usual[9,10] *Secondary care* Treatment as usual[11] Waitlist[12,13]	*Primary care* Treatment as usual[14,16,17,18,19,20] Enhanced treatment as usual[15] *Secondary care* Waitlist[3,21,26] Treatment as usual[22,23,24,25]
Notes: [1]Beeber 2010; [2]Schulberg 1996; [3]Lemmens 2015 /2016; [4]Swartz 2008; [5]Van Schaik 2006; [6]Scott 1992; [7]Ward 2000; [8]MacPherson 2013; [9]Dalgard 2006; [10]Ekers 2011; [11]Gawrysiak 2009; [12]McIndoo 2016; [13]Spek 2007; [14]Cramer 2011; [15]Dwight-Johnson 2011; [16]Laidlaw 2008; [17]Miranda 2003; [18]Scott 1992; [19]Serfaty 2009; [20]Ward 2000; [21]Kohtala 2015; [22]Losada 2015; [23]Mohr 2011; [24]Naeem 2015; [25]Scott 1997; [26]Selmi 1990 Lemmens 2015/2016, MacPherson 2013, McIndoo 2016, Miranda 2003, Schulberg 1996, Selmi 1990, Serfaty 2009, Spek 2007 and Ward 2000 are three-armed trials and Scott 1992 is a four-armed trial but where possible the demographics reported here are for only the two relevant arms.			

Table 26: Study information table for trials included in the sub-analysis of primary care versus secondary care (part 3 – self-help interventions)

	Self-help versus TAU/waitlist	Self-help with support versus TAU/waitlist
Total no. of studies (N randomised)	*Primary care* 3 (837) *Secondary care* 13 (1565)	*Primary care* 5 (1332) *Secondary care* 5 (393)
Study ID	*Primary care* Hallgren 2015[1] Joling 2011[2] Naylor 2010[3] *Secondary care* Geraedts 2014[4] Hoifodt 2013[5] Jamison 1995[6] Levin 2011[7] Liu 2009[8] Moldovan 2013[9] Moss 2012[10] Naeem 2014[11] Proudfoot 2004a[12] Salkovskis 2006[13] Scogin 1987[14] Selmi 1990[15] Spek 2007[16]	*Primary care* Gilbody 2015[17] Kessler 2009[18] Lovell 2008[19] Watkins 2012[20] Williams 2013c[21] *Secondary care* Choi 2012[22] Lamers 2015[23] Perini 2009[24] Ruwaard 2009[25] Titov 2015[26]
Country	*Primary care* Sweden[1] Netherlands[2]	*Primary care* UK[17,18,19,20,21] *Secondary care*

	Self-help versus TAU/waitlist	Self-help with support versus TAU/waitlist
	US[3] *Secondary care* Netherlands[4,16] Norway[5] US[6,7,10,14,15] Taiwan[8] Romania[9] Pakistan[11] UK[12,13]	Australia[22,24,26] Netherlands[23,25]
Baseline depression severity	*Primary care* Less severe[1,2,3] *Secondary care* Less severe[4,5,6,8,9,10,11,12,14,15,16] More severe[7,13]	*Primary care* Less severe[17,19,20] More severe[18,21] *Secondary care* Less severe[22,23,25,26] More severe[24]
Age (mean)	*Primary care* NR by arm (43.0 for all 3 arms) [1] 81.5[2] 51.5[3] *Secondary care* 43.4[4] 36.0[5] 38.0[6] 43.5[7,12] 26.4[8] 22.2[9] 77.5[10] 33.5[11] 39.7[13] 71.3[14] 29.9[15] 55[16]	*Primary care* 39.9[17] 35.0[18] 37.6[19] 46.4[20] 41.8[21] *Secondary care* 39[22] 56.9[23] 49.3[24] 42[25] 65.3[26]
Sex (% female)	*Primary care* NR by arm (73% for all 3 arms) [1] 74[2] 84[3] *Secondary care* 62[4] 73[5,8] 84[6] 77[7,10] 90[9] 55.7[11] 74[12] 81[13] 76[14] 64[15] 63[16]	*Primary care* 67[17] 68[18,21] 74[19] 60[20] *Secondary care* 80[22] 77[23] 78[24] 69[25] 70[26]
Ethnicity (% BME)	*Primary care* NR[1,2]	*Primary care* NR[17,18,21]

	Self-help versus TAU/waitlist	Self-help with support versus TAU/waitlist
	8[3] *Secondary care* NR[4,5,9,11,13,14,16] 15[6] 10[7] 100[8] 19[10] 11[12] 0[15]	7[19] 0[20] *Secondary care* 100[22] NR[23,24,25,26]
Intervention	*Primary care* Computerised-CBT (CCBT) [1] Cognitive bibliotherapy[2,3] *Secondary care* Computerised-CBT (CCBT) [4,5,7,12,15,16] Cognitive bibliotherapy[6,8,9,10,11,13,14]	*Primary care* Computerised-CBT (CCBT) with support[17,18] Cognitive bibliotherapy with support[19,21] Cognitive bias modification with support[20] *Secondary care* Computerised-CBT (CCBT) with support[22,24,25,26] Cognitive bibliotherapy with support[23]
Comparison	*Primary care* Treatment as usual[1,2,3] *Secondary care* Treatment as usual[4,7,11,12,13] Waitlist[5,6,8,9,10,14,15,16]	*Primary care* Treatment as usual[17,19,20,21] Waitlist[18] *Secondary care* Waitlist[22,23,24,25,26]
Notes: [1]Hallgren 2015; [2]Joling 2011; [3]Naylor 2010; [4]Geraedts 2014; [5]Hoifodt 2013; [6]Jamison 1995; [7]Levin 2011; [8]Liu 2009; [9]Moldovan 2013; [10]Moss 2012; [11]Naeem 2014; [12]Proudfoot 2004a; [13]Salkovskis 2006; [14]Scogin 1987; [15]Selmi 1990; [16]Spek 2007; [17]Gilbody 2015; [18]Kessler 2009; [19]Lovell 2008; [20]Watkins 2012; [21]Williams 2013c; [22]Choi 2012; [23]Lamers 2015; [24]Perini 2009; [25]Ruwaard 2009; [26]Titov 2015 Hallgren 2015, Selmi 1990 and Spek 2007 are three-armed trials and Moldovan 2013 is a four-armed trial but where possible the demographics reported here are for only the two relevant arms.		

Update 2017

5.4.1.21 Crisis resolution and home treatment teams

Crisis resolution and home treatment teams include any type of crisis-oriented treatment of an acute psychiatric episode by staff with a specific remit to deal with such situations, in and beyond 'office hours'. This form of service aims to offer intensive home-based support in order to provide the best care for someone where this is the most appropriate setting. Traditionally, a depressive episode marked by serious risk to self (most often suicidal ideation and intent) or very severe deterioration to care for the self is managed by admission to an acute inpatient unit. However there is growing interest in attempting to manage episodes in the community. If done safely, it may avoid the stigma and costs associated with hospital admission. The evidence required to examine the benefits and harms associated with crisis resolution and home treatment teams would require trials that randomise participants to crisis-intervention care versus standard (inpatient) care. However, the large majority of patients with depression are never admitted to hospital, meaning that there is limited evidence from RCTs to determine the value of crisis resolution teams for depression-specific populations. Indeed, no RCT evidence was identified that specifically addressed this setting for adults with depression. The GC therefore agreed to consider a wider evidence base including non-psychotic severe mental illness and a wider definition of important but not critical outcomes (including non-depression-specific measures of psychological functioning and satisfaction). A systematic review (Murphy 2015; updated version of Joy 2003 used in

2009 guideline) was identified that examined crisis intervention for people with severe mental illness. This Cochrane review was used as a source of studies with inclusion criteria into this review of over 50% of the population having a non-psychotic disorder.

Of the eight RCTs included in Murphy 2015, one of these studies met the >50% non-psychotic disorder inclusion criterion (Johnson 2005), see Table 27 for study characteristics.

Evidence for this comparison is summarised in the clinical GRADE evidence profile below (Table 28). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J1.2.

Table 27: Study information table for trials included in the analysis of crisis resolution and home treatment care versus standard care for adults with non-psychotic severe mental illness

	Crisis resolution team care versus standard care
Total no. of studies (N randomised)	1 (260)
Study ID	Johnson 2005
Country	UK
Diagnosis	25% Schizophrenia or schizoaffective disorder; 10% Bipolar affective disorder; 7% Other psychosis; 30% Unipolar depression; 13% Personality disorder; 4% Other non-psychotic disorder; 5% Substance misuse only (data only reported for 123/135 of experimental group so percentages do not add up to 100%)
Age range (mean)	NR (37.9)
Sex (% female)	49
Ethnicity (% BME)	22
Intervention	Crisis resolution team augmented existing acute services and aimed to assess all patients and manage them at home if feasible. Staff were available 24 hours but on call from home after 10pm
Comparison	Standard care included care from the inpatient unit, crisis houses, and community mental health teams
Duration of follow-up	6 months (outcomes also assessed at 8 weeks)

Table 28: Summary of findings table for crisis resolution and home treatment care compared to standard care for adults with non-psychotic severe mental illness

Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Standard care	**Crisis resolution team care**				
Lost to follow-up Number of participants lost to follow-up by the end of the study Follow-up: mean 12 months	Study population		**RR 0.93** (0.49 to 1.73)	260 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	136 per 1000	**126 per 1000** (67 to 235)				
	Moderate					
	136 per 1000	**126 per 1000** (67 to 235)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Standard care	Crisis resolution team care				
Symptom severity (BPRS) Brief Psychiatric Rating Scale (BPRS) 8 weeks after crisis Follow-up: mean 8 weeks		The mean symptom severity (BPRS) in the intervention groups was 0.29 standard deviations lower (0.56 to 0.02 lower)		211 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]	SMD -0.29 (-0.56 to -0.02)
Admission as inpatient Number of participants that had been admitted to a psychiatric ward within 6 months after crisis Follow-up: mean 6 months	**Study population** **677 per 1000**	**291 per 1000** (217 to 386)	**RR 0.43** (0.32 to 0.57)	258 (1 study)	⊕⊖⊖⊖ **very low**[1,2,5]	
	Moderate **677 per 1000**	**291 per 1000** (217 to 386)				
Bed days in hospital Number of bed days in hospital for those admitted within 6 months after crisis Follow-up: mean 6 months		The mean bed days in hospital in the intervention groups was **18.9 lower** (29.38 to 8.42 lower)		257 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]	
Satisfaction Client Satisfaction Questionnaire - 8 item version (CSQ-8) 8 weeks after crisis Follow-up: mean 8 weeks		The mean satisfaction in the intervention groups was **0.23 standard deviations higher** (0.03 lower to 0.49 higher)		226 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]	SMD 0.23 (-0.03 to 0.49)
Quality of life Manchester short assessment of quality of life (MANSA) 8 weeks after crisis Follow-up: mean 8 weeks		The mean quality of life in the intervention groups was **0.11 standard deviations lower** (0.37 lower to 0.16 higher)		217 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]	SMD -0.11 (-0.37 to 0.16)
Social functioning (8 weeks after crisis) Life Skills Profile (LSP)		The mean social functioning (8 weeks after crisis) in the intervention groups was **0.2 standard**		257 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]	SMD 0.2 (-0.05 to 0.44)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Standard care	Crisis resolution team care				
Follow-up: mean 8 weeks		**deviations higher** (0.05 lower to 0.44 higher)				
Social functioning (at endpoint) Life Skills Profile (LSP) Follow-up: mean 6 months		The mean social functioning (at endpoint) in the intervention groups was **0.06 standard deviations higher** (0.18 lower to 0.31 higher)		255 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]	SMD 0.06 (-0.18 to 0.31)

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups and baseline and non-blind participants, intervention administrator(s) and outcome assessor(s)
[2] Not depression-specific population
[3] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[4] N<400
[5] Events<300

Update 2017

5.4.1.31 Inpatient care

No RCT evidence was identified that specifically addressed this setting. Therefore the GC considered indirect evidence in the form of sub-analyses of the NMA dataset (acute treatment of depressive episodes). In fact, a comparison of inpatient and outpatient settings was an a priori sub-analysis of the NMA dataset (for study characteristics see Chapter 7).

Sufficient data (2 or more RCTs per comparison) were only available to conduct a subgroup analysis of inpatient compared with outpatient care for one comparison, exercise versus attention placebo/TAU.

No statistically significant subgroup differences were found between inpatient and outpatient populations for exercise versus attention-placebo or treatment as usual (Depression symptomatology: $Chi^2 = 0.05$, df = 1, p = 0.82; Discontinuation for any reason: $Chi^2 = 1.80$, df = 1, p = 0.18).

5.4.1.43 Acute psychiatric day hospital care

Acute psychiatric day hospitals are units that provide diagnostic and treatment services for acutely ill individuals who would otherwise be treated in traditional psychiatric inpatient units. Two studies were identified that specifically addressed this setting for adults with depression, however, only 1 of these was an RCT and could be included (Dinger 2014). The GC therefore agreed to consider a wider evidence base including non-psychotic severe mental illness and a wider definition of important but not critical outcomes (including satisfaction, social functioning, carer distress and non-depression-specific measures of psychological functioning). A systematic review (Marshall 2011) was identified that compared day hospital to inpatient care for people with acute psychiatric disorders. This Cochrane review was used as a source of studies with inclusion criteria into this review of over 50% of the population having a non-psychotic disorder.

Of the ten RCTs included in Marshall 2011, 5 of these studies met the >50% non-psychotic disorder inclusion criterion (Creed 1990; Creed 1997; Dick 1985; Kallert 2007; Schene 1993), see Table 29 for study characteristics.

Evidence for this comparison is summarised in the clinical GRADE evidence profile below (Table 30). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J1.2.

Table 29: Study information table for trials included in the meta-analysis of acute day hospital care versus inpatient care for adults with non-psychotic severe mental illness

	Acute day hospital care versus inpatient care
Total no. of studies (N randomised)	6 (1763)
Study ID	Creed 1990[1] Creed 1997[2] Dick 1985[3] Dinger 2014[4] Kallert 2007[5] Schene 1993[6]
Country	UK[1,2,3] Germany[4] Germany, UK, Poland, Slovakia and Czech Republic[5] Netherlands[6]
Diagnosis	27% Schizophrenia; 20% Depression; 9% Mania; 27% Neurotic disorder; 9% Personality disorder; 8% Addiction/organic disorder[1] 43% Schizophrenia; 34% Depression; 23% Neurosis[2] Neurosis (56% depressive neurosis), personality disorder, or adjustment reaction[3] 97.7% had a major depressive episode, 2.3% had primary dysthymia[4] 27% Schizophrenia, schizotypal, delusional, and other non-mood psychotic disorders (ICD-10 F20-F29); 41% Mood [affective] disorders (ICD-10 F30-F39); 22% Anxiety, dissociative, stress-related, somatoform and other nonpsychotic mental disorders (ICD-10 F40-F49); 9% Disorders of adult personality and behaviour (ICD-10 F60-F69)[5] 21% Psychosis; 38% Mood disorders; 24% Anxiety disorders; 10% Eating disorders; 8% Other[6]
Age range (mean)	Range NR (42.5)[1] Range NR (38.0)[2] Range NR (~ 35)[3] 18–55 (35.1)[4] Range NR (~ 38)[5] Range NR (31.9)[6]
Sex (% female)	51[1] 43[2] 68[3] 50[4] 56[5] 58[6]
Ethnicity (% BME)	NR[1,3,4,5,6] 18[2]

	Acute day hospital care versus inpatient care
Intervention	Acute day hospital care. Teaching hospital serving small socially deprived inner city area. Day hospital designed to take acute admissions because of few beds (8 nurses, 3 OTs) [1] Acute day hospital care. Teaching hospital serving small socially deprived inner city area. Day hospital designed to take acute admissions because of few beds (CPN out of hours) [2] Acute day hospital care. 2 trained staff + OT, patient/staff ratio: 12.5:1, individual counselling, groups, activities and medication[3] Acute day hospital care. Therapeutic staff were the same for both treatment arms. Both groups received equal amounts of psychotherapeutic interventions. Day-clinic patients attended therapy on 5 weekdays from 8 a.m. to 4 p.m. (8 weeks of treatment) [4] Acute day hospital care. Provided between 15 and 35 places, mean staff hours per week per treatment place ranged from 8.8 to 16.0. Staff patient ratios not reported[5] Acute day hospital care. Provided 24 places. For each day treatment patient, a 0.08 full-time equivalent social psychiatric nurse was available[6]
Comparison	Inpatient care (routine inpatient)[1,2,5] Inpatient care. Mixed sex and female wards[3] Inpatient care. Therapeutic staff were the same for both treatment arms. Both groups received equal amounts of psychotherapeutic interventions. Inpatients were free to leave the unit outside of night hours and therapy sessions and spent 6 weekends at home (8 weeks of treatment) [4] Inpatient care. Open inpatient ward with 20 beds. For each inpatient, a 0.40 full-time equivalent psychiatric nurse was available[6]
Duration of follow-up	12 months[1,2,3] 3 months[4] 14 months[5] 13 months[6]
Notes: [1]Creed 1990; [2]Creed 1997; [3]Dick 1985; [4]Dinger 2014; [5]Kallert 2007; [6]Schene 1993	

Table 30: Summary of findings table for acute day hospital care compared to inpatient care for adults with non-psychotic severe mental illness

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Inpatient care	Acute day hospital care				
Lost to follow-up Number of participants lost to follow-up by the end of the study Follow-up: 3-14 months	**Study population**		**RR 1.25** (0.96 to 1.63)	1763 (6 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	315 per 1000	**394 per 1000** (303 to 514)				
	Moderate					
	178 per 1000	**222 per 1000** (171 to 290)				
	Study population					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Inpatient care	Acute day hospital care				
Death (suicide) Number of participants that committed suicide during the study period Follow-up: mean 14 months	**6 per 1000**	**1 per 1000** (0 to 14)	**RR 0.12** (0.01 to 2.41)	1117 (1 study)	⊕⊖⊖⊖ **very low**[4,5]	
	Moderate					
	6 per 1000	**1 per 1000** (0 to 14)				
Remission of psychiatric symptoms Present State Examination: Index of Definition≤4/<7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 3-13 months	**Study population**		**RR 0.91** (0.65 to 1.26)	151 (2 studies)	⊕⊖⊖⊖ **very low**[2,6,7,8]	
	465 per 1000	**423 per 1000** (302 to 586)				
	Moderate					
	369 per 1000	**336 per 1000** (240 to 465)				
Response Number of people showing ≥47% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 3 months	**Study population**		**RR 0.62** (0.26 to 1.5)	44 (1 study)	⊕⊖⊖⊖ **very low**[7,9,10]	
	400 per 1000	**248 per 1000** (104 to 600)				
	Moderate					
	400 per 1000	**248 per 1000** (104 to 600)				
Symptom severity (2-3 months post-admission) Comprehensive Psychopathological Rating Scale (CPRS; change score)/Brief Psychiatric Rating Scale (BPRS; change score)/Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: 2-3 months		The mean symptom severity (2-3 months post-admission) in the intervention groups was **0.05 standard deviations higher** (0.22 lower to 0.33 higher)		1281 (3 studies)	⊕⊖⊖⊖ **very low**[2,11,12]	SMD 0.05 (-0.22 to 0.33)
Symptom severity (12-14 months post-admission) Comprehensive Psychopathological Rating Scale (CPRS; change score)/Brief Psychiatric Rating Scale (BPRS; change score) Follow-up: 12-14 months		The mean symptom severity (12-14 months post-admission) in the intervention groups was **0.19 standard deviations lower** (0.81 lower to 0.42 higher)		1249 (2 studies)	⊕⊖⊖⊖ **very low**[2,11,13,14]	SMD -0.19 (-0.81 to 0.42)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Inpatient care	Acute day hospital care				
Duration of index admission Number of days/months in hospital Follow-up: 12-14 months		The mean duration of index admission in the intervention groups was **0.55 standard deviations higher** (0.44 to 0.65 higher)		1535 (4 studies)	⊕⊖⊖⊖ **very low**[2,11]	SMD 0.55 (0.44 to 0.65)
Readmission Number of patients readmitted to hospital Follow-up: mean 12 months	**Study population**		**RR 0.79** (0.41 to 1.52)	372 (3 studies)	⊕⊖⊖⊖ **very low**[2,5,8,12,15]	
	249 per 1000	**196 per 1000** (102 to 378)				
	Moderate					
	215 per 1000	**170 per 1000** (88 to 327)				
Discharge Number of participants discharged from hospital within 3 months of admission Follow-up: mean 3 months	**Study population**		**RR 0.6** (0.4 to 0.91)	89 (1 study)	⊕⊖⊖⊖ **very low**[2,8,15,16]	
	688 per 1000	**412 per 1000** (275 to 626)				
	Moderate					
	688 per 1000	**413 per 1000** (275 to 626)				
Service utilisation: Emergency contacts Number of participants making emergency contacts within 4 months post-admission Follow-up: mean 4 months	**Study population**		**RR 2.37** (0.98 to 5.71)	83 (1 study)	⊕⊖⊖⊖ **very low**[2,3,8,17]	
	133 per 1000	**316 per 1000** (131 to 761)				
	Moderate					
	133 per 1000	**315 per 1000** (130 to 759)				
Service utilisation: Outpatient contact Number of participants making outpatient contacts within 4 months post-admission Follow-up: mean 4 months	**Study population**		**RR 1.38** (0.73 to 2.62)	83 (1 study)	⊕⊖⊖⊖ **very low**[2,5,8,17]	
	267 per 1000	**368 per 1000** (195 to 699)				
	Moderate					
	267 per 1000	**368 per 1000** (195 to 700)				

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Inpatient care	Acute day hospital care				
Satisfaction Number of participants satisfied or very satisfied with their treatment Follow-up: mean 4 months	**Study population**		**RR 1.93** (1.33 to 2.81)	83 (1 study)	⊕⊖⊖⊖ **very low**[2,8,16,17]	
	422 per 1000	**815 per 1000** (562 to 1000)				
	Moderate					
	422 per 1000	**814 per 1000** (561 to 1000)				
Satisfaction Cliet Assessment of Treatment (CAT) Follow-up: mean 2 months		The mean satisfaction in the intervention groups was **0.03 standard deviations higher** (0.09 lower to 0.15 higher)		1117 (1 study)	⊕⊖⊖⊖ **very low**[2,11]	SMD 0.03 (-0.09 to 0.15)
Quality of life (2-months post-admission) Manchester short assessment of quality of life (MANSA) Follow-up: mean 2 months		The mean quality of life (2-months post-admission) in the intervention groups was **0.01 standard deviations higher** (0.11 lower to 0.13 higher)		1117 (1 study)	⊕⊖⊖⊖ **very low**[2,11]	SMD 0.01 (-0.11 to 0.13)
Quality of life (14-months post-admission) Manchester short assessment of quality of life (MANSA) Follow-up: mean 14 months		The mean quality of life (14-months post-admission) in the intervention groups was **0.01 standard deviations higher** (0.11 lower to 0.13 higher)		1117 (1 study)	⊕⊖⊖⊖ **very low**[2,11]	SMD 0.01 (-0.11 to 0.13)
Social functioning response 2 role disabilities or less on Groningen Social Disabilities Schedule (GSDS)/Number of participants living in the community and social functioning at previous level (according to the social performance and	**Study population**		**RR 1.36** (0.94 to 1.96)	181 (2 studies)	⊕⊖⊖⊖ **very low**[2,8,18,19]	
	333 per 1000	**453 per 1000** (313 to 653)				
	Moderate					
	342 per 1000	**465 per 1000** (321 to 670)				

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Inpatient care	Acute day hospital care				
behaviour assessment schedule) Follow-up: 12-13 months						
Social functioning impairment (2-months post-admission) Groningen Social Disabilities Schedule, Second revision (GSDS-II) Follow-up: mean 2 months		The mean social functioning impairment (2-months post-admission) in the intervention groups was **0.3 standard deviations lower** (0.42 to 0.19 lower)		1117 (1 study)	⊕⊖⊖⊖ **very low**[2,11]	SMD -0.3 (-0.42 to -0.19)
Social functioning impairment (14-months post-admission) Groningen Social Disabilities Schedule, Second revision (GSDS-II) Follow-up: mean 14 months		The mean social functioning impairment (14-months post-admission) in the intervention groups was **0.15 standard deviations lower** (0.27 to 0.04 lower)		1117 (1 study)	⊕⊖⊖⊖ **very low**[2,11]	SMD -0.15 (-0.27 to -0.04)
Carer distress (3-months post-admission) General Health Questionnaire (GHQ; change score) Follow-up: mean 3 months		The mean carer distress (3-months post-admission) in the intervention groups was **1.1 lower** (3.15 lower to 0.95 higher)		77 (1 study)	⊕⊖⊖⊖ **very low**[2,14,15]	
Carer distress (12-months post-admission) General Health Questionnaire (GHQ; change score) Follow-up: mean 12 months		The mean carer distress (12-months post-admission) in the intervention groups was **0.4 lower** (2.98 lower to 2.18 higher)		55 (1 study)	⊕⊖⊖⊖ **very low**[2,14,15]	

Notes:
[1] Randomisation method was unclear (or high risk associated with it due to significant baseline differences). Non-blind participants, intervention administrator(s) and unclear blinding of, or non-blind, outcome assessor(s)
[2] Non depression-specific population
[3] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Inpatient care	Acute day hospital care				

[4] High risk of bias associated with randomisation method due to significant difference between groups at baseline. Non-blind participants, intervention administrator(s) and outcome assessor(s). Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[5] 95% CI crosses line of no effect and both threshold for clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[6] Unclear randomisation method and method of allocation concealment. Non-blind participants and intervention administrator(s) and unclear blinding of outcome assessment
[7] 95% CI crosses line of no effect and threshold for clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[8] Data cannot be extracted for all outcomes (measure of variance not reported)
[9] Unclear blinding of allocation concealment. Non-blind participants and intervention administrator(s) and unclear blinding of outcome assessment. Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[10] A non-standard definition of response selected (e.g. 47% rather than 50%)
[11] High risk of bias associated with randomisation method due to significant difference between groups at baseline. Non-blind participants, intervention administrator(s) and outcome assessment. Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[12] I-squared>50%
[13] I-squared>80%
[14] 95% CI crosses both line of no effect and threshold for clinically important benefit (SMD -0.5)
[15] Non-blind participants, intervention administrator(s) and outcome assessment
[16] Events<300
[17] Unclear randomisation method and allocation concealment, and non-blind participants, intervention administrator(s) and outcome assessment
[18] Non-blind participants and intervention administrator(s) and non-blind, or unclear blinding of, outcome assessment. Unclear risk of attrition bias (drop-out>20% but difference between groups<20%)
[19] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)

Update 2017

5.4.1.51 Non-acute day hospital care and recovery centres

Although the earliest use of day hospitals in mental healthcare was to provide an alternative to inpatient care (Cameron, 1947), non-acute day hospitals, psychiatric day hospitals offering continuing care, have also been used for people with refractory mental health problems unresponsive to treatment in outpatient clinics and may include patients with depressive disorders who have residual or persistent symptoms. No RCT evidence was identified that specifically addressed this setting for adults with depression. The GC therefore agreed to consider a wider evidence base including non-psychotic severe mental illness and a wider definition of important but not critical outcomes (including non-depression-specific measures of psychological functioning and satisfaction). A systematic review (Marshall 2001) was identified that examined the use of day hospitals as an alternative to outpatient care for people with psychiatric disorders. This Cochrane review was used as a source of studies with inclusion criteria into this review of over 50% of the population having a non-psychotic disorder.

Of the eight studies included in Marshall 2001, three of these studies met the >50% non-psychotic disorder inclusion criterion (Dick 1991; Glick 1986; Tyrer 1979), see Table 31 for study characteristics.

Evidence for this comparison is summarised in the clinical GRADE evidence profile below (Table 32). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J1.2.

Table 31: Study information table for trials included in the meta-analysis of non-acute day hospital care versus outpatient care for adults with non-psychotic severe mental illness

	Non-acute day hospital care versus outpatient care
Total no. of studies (N randomised)	3 (281)
Study ID	Dick 1991[1] Glick 1986[2] Tyrer 1979[3]
Country	UK[1,3] US[2]
Diagnosis	92% DSM-III major depressive disorder; 8% dysthymic disorder[1] 47% Schizophrenia; 53% Major affective disorder[2] Neurotic disorder (severe enough for day hospital treatment) [3]
Age range (mean)	NR (52% <45 years) [1] Range NR (35) [2] 16-60 years (mean NR) [3]
Sex (% female)	75[1] 63[2] NR[3]
Ethnicity (% BME)	NR
Intervention	Non-acute day hospital care. Places for up to 40 patients. Treatment is eclectic, with a focus on time structuring and socialisation, and a problem-orientated supportive/behavioural rather than a psychodynamic approach. Staffing comprises three sessions per week of consultant time, three sessions per week of support medical time, three full-time trained nurses, and one full-time occupational therapist. Mean duration of day treatment was 10.7 weeks[1] Non-acute day hospital care. Transitional day care following inpatient admission (about 15 hours/week and limited to 6-12 weeks) involving milieu, family, supportive & group therapy, medication, care management, recreation & dance therapy, and discharge planning[2] Non-acute day hospital care. Two different types of day hospital: one specialising in neurotic disorders (well-staffed with psychotherapeutic orientation) and the other a standard day hospital (psychiatrists, nurses, occupational & art therapists) [3]
Comparison	Outpatient care. Patients allocated to continued outpatient treatment were seen approximately monthly and given advice on relaxation, anxiety management, and alternative approaches to time structuring and handling relationships[1] Outpatient care. Outpatient follow-up post-inpatient admission involving 6-12 weeks in outpatient group therapy (90 mins/week), medication management and 24 hour crisis intervention[2] Outpatient care (routine outpatient) [3]
Duration of follow-up	6 months[1] 12 months[2] 24 months[3]
Notes: [1]Dick 1991; [2]Glick 1986; [3]Tyrer 1979	

Table 32: Summary of findings table for non-acute day hospital care compared to outpatient care for adults with non-psychotic severe mental illness

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	Non-acute day hospital care versus outpatient care				
Lost to follow-up Number of participants lost to follow-up by the end of the study Follow-up: 6-24 months	**Study population**		**RR 0.81** (0.24 to 2.7)	281 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,3,4,5]	
	207 per 1000	**168 per 1000** (50 to 559)				
	Moderate					
	207 per 1000	**168 per 1000** (50 to 559)				
Death (all causes) Number of participants who died due to any causes during the study period Follow-up: mean 24 months	**Study population**		**RR 2.42** (0.23 to 25.85)	106 (1 study)	⊕⊖⊖⊖ **very low**[3,4,6]	
	17 per 1000	**42 per 1000** (4 to 446)				
	Moderate					
	17 per 1000	**41 per 1000** (4 to 439)				
Symptom severity (4-6 months post-admission) Psychiatric Evaluation Form (change score)/Present State Examination (change score) Follow-up: 4-6 months		The mean symptom severity (4-6 months post-admission) in the intervention groups was **0.08 standard deviations higher** (0.72 lower to 0.88 higher)		144 (2 studies)	⊕⊖⊖⊖ **very low**[3,7,8,9]	SMD 0.08 (-0.72 to 0.88)
Symptom severity (8-12 months post-admission) Psychiatric Evaluation Form (change score)/Present State Examination (change score) Follow-up: 8-12 months		The mean symptom severity (8-12 months post-admission) in the intervention groups was **0.15 standard deviations lower** (0.49 lower to 0.19 higher)		139 (2 studies)	⊕⊖⊖⊖ **very low**[3,7,10,11]	SMD -0.15 (-0.49 to 0.19)
Admission as inpatient Number of participants admitted	**Study population**		**RR 1.26** (0.52 to 3.06)	281 (3 studies)	⊕⊖⊖⊖ **very low**[3,4,12]	
	83 per 1000	**104 per 1000** (43 to 253)				

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	Non-acute day hospital care versus outpatient care				
into inpatient care during the study period Follow-up: 6-12 months	**Moderate**					
	80 per 1000	**101 per 1000** (42 to 245)				
Satisfaction Number of participants satisfied or very satisfied with their treatment Follow-up: 4-6 months	**Study population**		**RR 1** (0.47 to 2.12)	198 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,8,13]	
	632 per 1000	**632 per 1000** (297 to 1000)				
	Moderate					
	628 per 1000	**628 per 1000** (295 to 1000)				
Global functioning (6-months post-admission) Global Assessment Scale (GAS; change score) Follow-up: mean 6 months		The mean global functioning (6-months post-admission) in the intervention groups was **0.04 standard deviations higher** (0.53 lower to 0.61 higher)		52 (1 study)	⊕⊖⊖⊖ **very low**[3,9,14]	SMD 0.04 (-0.53 to 0.61)
Global functioning (12-months post-admission) Global Assessment Scale (GAS; change score) Follow-up: mean 12 months		The mean global functioning (12-months post-admission) in the intervention groups was **0.12 standard deviations lower** (0.7 lower to 0.45 higher)		51 (1 study)	⊕⊖⊖⊖ **very low**[3,14,15]	SMD -0.12 (-0.7 to 0.45)
Social functioning (4-6 months post-admission) Social Adjustment Scale-Self Report (SAS-SR; change score)/Social Functioning Scale (SFS; change score) Follow-up: 4-6 months		The mean social functioning (4-6 months post-admission) in the intervention groups was **0.2 standard deviations lower** (0.54 lower to 0.14 higher)		141 (2 studies)	⊕⊖⊖⊖ **very low**[3,7,11,15]	SMD -0.2 (-0.54 to 0.14)

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	Non-acute day hospital care versus outpatient care				
Social functioning (8-12 months post-admission) Social Adjustment Scale-Self Report (SAS-SR; change score)/Social Functioning Scale (SFS; change score) Follow-up: 8-12 months		The mean social functioning (8-12 months post-admission) in the intervention groups was **0.31 standard deviations lower** (0.65 lower to 0.03 higher)		140 (2 studies)	⊕⊖⊖⊖ **very low**[3,7,11,15]	SMD -0.31 (-0.65 to 0.03)

Notes:
[1] Unclear randomisation method and non-blind participants and intervention administrator(s)
[2] I-squared>50%
[3] Non-depression specific population
[4] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[5] Data cannot be extracted or is not reported for all outcomes
[6] Unclear randomisation method and non-blind participants and intervention administrator(s). Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[7] Unclear randomisation method and non-blind participants and intervention administrator(s). Risk of attrition bias is unclear or high (drop-out>20% and ITT analysis not used)
[8] I-squared>80%
[9] 95% CI crosses line of no effect and threshold for both clinically important benefit (SMD -0.5) and clinically important harm (SMD 0.5)
[10] N<400
[11] Data is not reported for longest follow-up
[12] Unclear randomisation method and method of allocation concealment. Non-blind participants and intervention administrator(s) and unclear blinding of outcome assessment. Unclear risk of attrition bias (drop-out>20%)
[13] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[14] Unclear randomisation method and method of allocation concealment. Non-blind participants and intervention administrator(s) and unclear blinding of outcome assessment. High risk of attrition bias as drop-out>20%, difference between groups>20% and completer analysis used
[15] 95% CI crosses both line of no effect and threshold for clinically important benefit (SMD-0.5)

5.4.1.61 Specialist tertiary affective disorders settings

2 One RCT (Morriss 2016) was found and included, that compares a specialist depression
3 service to usual specialist mental health care for adults with persistent depression, see Table
4 33 for study characteristics.

5 Evidence for this comparison is summarised in the clinical GRADE evidence profile below
6 (Table 34). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix
7 M and the full study characteristics, comparisons and outcomes tables in Appendix J1.2.

Table 33: Study information table for trials included in the analysis of specialist depression service versus usual specialist mental health care for adults with persistent depression

	Specialist depression service versus usual specialist mental health care
Total no. of studies (N randomised)	1 (187)
Study ID	Morriss 2016
Country	UK
Diagnosis	DSM-IV MDD (persistent depression defined as non-response to secondary mental health care for at least 6 months)
Age range (mean)	Range NR (46.5)
Sex (% female)	56
Ethnicity (% BME)	NR
Intervention	Specialist depression service. The specialist depression service offered NICE-recommended pharmacological (specialist pharmacotherapy for depression) and psychological (CBT) treatments for depression as a specialty service within specialist mental health care, with a collaborative care approach between psychiatrists and cognitive behavioural therapists over 12 months, followed by a graduated transfer of care up to an additional 3 months (15 months in total) to primary care or usual specialist mental health care
Comparison	Usual specialist mental health care. This treatment was directed by a consultant psychiatrist, usually from a community mental health team. It usually consisted of individual work by the psychiatrist in secondary care and treatment with pharmacotherapy (often by use of augmentation and change strategies for depression not responding to trials of single antidepressants), sometimes shared reviews with primary care, and sometimes used psychosocial interventions such as CBT, counselling, or community psychiatric nurse support. Little evidence was available of joint reviews of progress and regular professional meetings being used. Responses to medication and treatment were reviewed only by the provider of the treatment
Duration of follow-up	18 months

Table 34: Summary of findings table for specialist depression service compared to usual specialist mental health care for adults with persistent depression

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Usual specialist mental health care	Specialist depression service				
Lost to follow-up Number of participants lost to follow-up by the end of the study Follow-up: mean 18 months	**Study population**		**RR 0.68** (0.48 to 0.97)	187 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
	489 per 1000	**333 per 1000** (235 to 475)				
	Moderate					
	489 per 1000	**333 per 1000** (235 to 474)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Usual specialist mental health care	Specialist depression service				
Self-harm Number of participants who had episodes of self-harm during the study period Follow-up: mean 18 months	**Study population**		**RR 0.51** (0.05 to 5.48)	187 (1 study)	⊕⊖⊖⊖ **very low**[3,4]	
	21 per 1000	**11 per 1000** (1 to 117)				
	Moderate					
	21 per 1000	**11 per 1000** (1 to 115)				
Response Hamilton Rating Scale for Depression (HAM-D) - definition for response not reported Follow-up: mean 18 months	**Study population**		**RR 1.63** (1.05 to 2.51)	187 (1 study)	⊕⊕⊖⊖ **low**[2,3]	
	245 per 1000	**399 per 1000** (257 to 614)				
	Moderate					
	245 per 1000	**399 per 1000** (257 to 615)				
Remission Hamilton Rating Scale for Depression (HAM-D) - cut-off for remission not reported Follow-up: mean 18 months	**Study population**		**RR 2.02** (1.08 to 3.8)	187 (1 study)	⊕⊕⊖⊖ **low**[2,3]	
	128 per 1000	**258 per 1000** (138 to 485)				
	Moderate					
	128 per 1000	**259 per 1000** (138 to 486)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 18 months		The mean depression symptomatology in the intervention groups was **0.62 standard deviations lower** (0.92 to 0.33 lower)		187 (1 study)	⊕⊕⊖⊖ **low**[3,5]	SMD -0.62 (-0.92 to -0.33)
Global functioning Global Assessment of Functioning (GAF; change score) Follow-up: mean 18 months		The mean global functioning in the intervention groups was **0.49 standard deviations higher** (0.19 to 0.78 higher)		187 (1 study)	⊕⊕⊖⊖ **low**[3,5]	SMD 0.49 (0.19 to 0.78)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Usual specialist mental health care	Specialist depression service				
Social functioning Social Adjustment Scale-modified (SAS-M; change score) Follow-up: mean 18 months		The mean social functioning in the intervention groups was **0.46 standard deviations higher** (0.17 to 0.75 higher)		187 (1 study)	⊕⊕⊖⊖ **low**[3,5]	SMD 0.46 (0.17 to 0.75)

Notes:
[1] Non-blind participants and intervention administrator(s)
[2] Events<300
[3] Non-blind participants and intervention administrator(s). Risk of attrition bias is unclear (drop-out>20% but difference between groups<20% and ITT analysis used)
[4] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[5] N<400

Update 2017

5.4.1.71 Community mental health teams (CMHTs)

No RCT evidence was identified that specifically addressed this setting for adults with depression. The GC therefore agreed to consider a wider evidence base including non-psychotic severe mental illness and a wider definition of important but not critical outcomes (including non-depression-specific measures of psychological functioning and satisfaction). A systematic review (Malone 2007) was identified that examined community mental health teams (CMHTs) for people with severe mental illnesses and disordered personality. This Cochrane review was used as a source of studies with inclusion criteria into this review of over 50% of the population having a non-psychotic disorder.

Of the three studies included in Malone 2007, one of these studies met the >50% non-psychotic disorder inclusion criterion (Merson 1992), see Table 35 for study characteristics.

Evidence for this comparison is summarised in the clinical GRADE evidence profile below (Table 36). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J1.2.

Table 35: Study information table for trials included in the meta-analysis of community mental health teams (CMHTs) versus standard care for adults with non-psychotic severe mental illness

	Community mental health teams (CMHTs) versus standard care
Total no. of studies (N randomised)	1 (100)
Study ID	Merson 1992
Country	UK
Diagnosis	38% ICD-10 Schizophrenia and related disorders; 32% Mood disorder; 25% Neurotic and stress-related disorders; 4% Substance misuse; 1% Personality disorder only

	Community mental health teams (CMHTs) versus standard care
Age range (mean)	Range NR (median 32)
Sex (% female)	60
Ethnicity (% BME)	32
Intervention	Community mental health team (CMHT). Early intervention from a multidisciplinary community-based team, open referral, in-home assessments, collaboration maintained with already involved agencies, clinical decisions by team consensus
Comparison	Standard care included conventional hospital-based psychiatric services, usually outpatient clinic assessments with occasional home visits
Duration of follow-up	3 months

Table 36: Summary of findings table for community mental health teams (CMHTs) compared to standard care for adults with non-psychotic severe mental illness

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	Community mental health teams (CMHTs) versus standard care				
Lost to follow-up Number of participants lost to follow-up by the end of the study Follow-up: mean 3 months	**Study population**		**RR 1.24** (0.49 to 3.16)	100 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	135 per 1000	**167 per 1000** (66 to 425)				
	Moderate					
	135 per 1000	**167 per 1000** (66 to 427)				
Death (all causes) Number of participants who died due to any causes during the study period Follow-up: mean 3 months	**Study population**		**RR 0.54** (0.05 to 5.78)	100 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	38 per 1000	**21 per 1000** (2 to 222)				
	Moderate					
	39 per 1000	**21 per 1000** (2 to 225)				
Symptom severity Comprehensive Psychopathological Rating Scale (CPRS) at endpoint Follow-up: mean 3 months		The mean symptom severity in the intervention groups was **0.06 standard deviations lower** (0.45 lower to 0.33 higher)		100 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4,5]	SMD -0.06 (-0.45 to 0.33)
	Study population					

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	Community mental health teams (CMHTs) versus standard care				
Admission as inpatient Number of participants admitted into inpatient care during the study period Follow-up: mean 3 months	**308 per 1000** **Moderate** **308 per 1000**	**145 per 1000** (65 to 323) **145 per 1000** (65 to 323)	**RR 0.47** (0.21 to 1.05)	100 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4,6]	
Admission as inpatient for >10 days Number of participants admitted into inpatient care for more than 10 days during the study period Follow-up: mean 3 months	**Study population** **212 per 1000** **Moderate** **212 per 1000**	**42 per 1000** (11 to 178) **42 per 1000** (11 to 178)	**RR 0.2** (0.05 to 0.84)	100 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4,7]	
Satisfaction Number of participants satisfied with their treatment Follow-up: mean 3 months	**Study population** **543 per 1000** **Moderate** **544 per 1000**	**832 per 1000** (614 to 1000) **832 per 1000** (615 to 1000)	**RR 1.53** (1.13 to 2.06)	87 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4,5]	
Satisfaction Service Satisfaction Score Follow-up: mean 3 months		The mean satisfaction in the intervention groups was **0.85 standard deviations higher** (0.41 to 1.29 higher)		87 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4,5]	SMD 0.85 (0.41 to 1.29)

Notes:
[1] Unclear randomisation method and non-blind participants and intervention administrator(s)
[2] Non-depression specific population
[3] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[4] Data cannot be extracted for all outcomes (no measure of variance reported)
[5] N<400
[6] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 0.75)
[7] Events<300

5.4.1.8 Residential services

No RCT or systematic review evidence was identified for residential service settings for adults with depression.

5.4.2 Economic evidence

No economic evidence on different settings for the delivery of care in adults with depression was identified by the systematic literature search. Details on the methods used for the systematic review of the economic literature are described in Chapter 3.

5.4.3 Clinical evidence statements

- Sub-analyses of NMA data suggests no significant differences between primary care and secondary care for amitriptyline compared to placebo, behavioural therapies compared to treatment as usual/waitlist, or cognitive and cognitive behavioural therapies compared to treatment as usual/waitlist for the acute treatment of depression in adults. There is some evidence for larger benefits of self-help (without support) in secondary care relative to primary care, however, there are also more secondary care studies and higher heterogeneity. The only other statistically significant difference between primary and secondary care is for self-help with support, with no differences in drop-out between self-help with support and treatment as usual/waitlist observed in primary care studies, however, in secondary care studies drop-out is significantly greater (over twice as high) in the self-help with support arm relative to treatment as usual or waitlist, suggesting there may be more issues with the acceptability of self-help with support in secondary care compared to in primary care.
- Very low quality single-RCT evidence (N=211-258) suggests a small but statistically significant benefit of crisis resolution team care relative to standard care on psychiatric symptom severity and service utilisation measures, including admission as an inpatient and bed days in hospital, for adults with non-psychotic severe mental illness. There is also a trend for a benefit in terms of patient satisfaction. However, evidence from the same RCT (N=217-260) suggests neither clinically important nor statistically significant benefits of crisis resolution team care on quality of life, social functioning or on acceptability or feasibility of the intervention (as measured by loss to follow-up).
- Sub-analyses of NMA data revealed no significant differences between inpatient and outpatient care for exercise compared to attention-placebo or treatment as usual for the acute treatment of depression in adults. Insufficient data is available to compare inpatient and outpatient care for any other comparison.
- Very low quality single-RCT (N= 83-89) evidence suggests a clinically important and statistically significant benefit of acute day hospital care relative to inpatient care on the number of adults with non-psychotic severe mental illness who are discharged within 3 months of admission and the number of people who are satisfied or very satisfied with their treatment. Very low quality evidence from another single-RCT (N= 1117) suggests a clinically important but not statistically significant benefit of acute day hospital care relative to inpatient care on the number of deaths due to suicide, a small but statistically significant benefit on a continuous measure of social functioning, and very low quality evidence from 2 RCTs (N=181) suggests a clinically important but not statistically significant benefit of acute day hospital care on a dichotomous measure of social functioning (the number of participants achieving significant improvement in social functioning). However, very low quality evidence from a single-RCT (N=44) including only adults with depression suggests a clinically important but not statistically significant benefit in favour of inpatient relative to acute day hospital care on the rate of response. In addition, very low quality evidence from 4 studies (N= 1535) suggests that adults with non-psychotic severe mental illness receiving acute day hospital care have a longer duration of index admission than those receiving inpatient care (clinically important and statistically significant). While very low quality evidence from a single-RCT (N=83) and from 6 RCTs (N=1763) suggests a

Update 2017

clinically important but not statistically significant harm of acute day hospital relative to inpatient care in terms of service utilisation measures (including emergency contacts and outpatient contact) and acceptability respectively. Very low quality evidence from 1-3 RCTs (N=151-1281) suggests neither clinically important nor statistically significant effects of acute day hospital care on the rate of remission, psychiatric symptom severity, readmission, a continuous measure of patient satisfaction, quality of life or carer distress.

- Very low quality evidence from 1-3 RCTs (N=51-281) suggests neither a clinically important nor statistically significant benefit of non-acute day hospital care relative to outpatient care on acceptability (as measured by the number of participants lost to follow-up), psychiatric symptom severity, satisfaction, global functioning, or social functioning, for adults with non-psychotic severe mental illness. While very low quality evidence from 1-3 RCTs (N=106-281) suggests clinically important but not statistically significant harms associated with non-acute day hospital care relative to outpatient care on the number of deaths (all causes) and the number of people admitted as an inpatient.
- Low quality single-RCT evidence (N=187) suggests clinically important and statistically significant benefits of a specialist depression service relative to usual specialist mental health care on the rate of response, the rate of remission, depression symptomatology, global functioning, social functioning and acceptability (as measured by the number of participants lost to follow-up) for adults with persistent depression. While, very low quality evidence from this same study suggests a clinically important but not statistically significant benefit of a specialist depression service on the number of participants who had episodes of self-harm during the study.
- Very low quality single-RCT evidence (N=87-100) suggests clinically important but not statistically significant benefits of community mental health team (CMHT) care relative to standard care on the number of deaths (all causes) and the number of participants admitted to inpatient care for adults with non-psychotic severe mental illness, and both clinically important and statistically significant benefits on the number of participants admitted to inpatient care for longer than 10 days, and both continuous and dichotomous measures of satisfaction. However, evidence from this same study suggests neither clinically important nor statistically significant benefits of CMHTs on psychiatric symptom severity or acceptability (as measured by the number of participants lost to follow-up).
- No evidence was identified for residential services for adults with depression.

Update 2017

5.4.4 Economic evidence statements

No economic evidence on different settings for the delivery of care in adults with depression is available.

5.4.5 From evidence to recommendations

5.4.5.1 Relative values of different outcomes

The GC identified depression symptomology, response, remission, relapse and acceptability (loss to follow-up) as the critical outcomes for this question. However, the GC also considered as important (but not critical) outcomes, service utilisation, satisfaction, social and global functioning and quality of life.

5.4.5.2 Trade-off between clinical benefits and harms

The best evidence to examine the benefits and harms associated with crisis resolution and home treatment teams would require trials that randomise participants to crisis-intervention care versus standard (inpatient) care. However, the large majority of patients with depression are never admitted to hospital, meaning that there is limited evidence from RCTs to determine the value of crisis resolution teams for depression-specific populations. The GC

therefore agreed to consider a wider evidence base including evidence on the care of people with severe mental illness.

Crisis resolution and home treatment team care appeared to improve psychiatric symptom severity and reduce inpatient admissions and time spent in hospital for adults with non-psychotic severe mental illness. However, the evidence came from a single study and was indirect, leading the GC to agree that a 'consider' rather than 'offer' recommendation was appropriate.

The GC recognised the potential benefits that crisis resolution and home treatment team care may bring to adults with severe depression, particularly those at significant risk of harming themselves through suicide attempts or self-neglect, in providing an alternative to inpatient treatment and thus potentially avoiding the stigma and costs associated with hospital admission. However, drawing on their clinical knowledge and expertise, the GC recognised that inpatient care was still an option for people with more severe depression who could not be adequately supported by a crisis resolution and home treatment team, particularly if they were socially isolated. They also recognised that crisis resolution and home treatment team care may have an important role in supporting people at home after an inpatient stay and so facilitate an early discharge, reducing the likelihood of a re-admission to hospital.

The GC also raised the importance of equity of access to interventions in inpatient care that is equivalent to those available in community settings. They therefore recommended that the full range of psychological interventions available in community settings should also be available in inpatient settings. They also recognised that the intensity and/or duration of these interventions may need to be altered commensurate with the level of severity and need in inpatient settings.

The GC considered the evidence for tertiary depression services and although they agreed that this may have advantages in improving treatments for a group of people with severe and complex depression who are not well served by current services, concerns were raised about the quality of the single-study evidence, particularly given the potential resource implications. Thus in the absence of further clinical and cost-effectiveness evidence the GC did not make a recommendation.

Update 2017

5.4.5.31 Trade-off between net health benefits and resource use

The GC considered the costs associated with crisis and intensive home treatment and estimated that these are higher than routine primary care but significantly lower than inpatient care. The GC expressed the opinion that, compared with routine primary care, crisis and intensive home treatment is often more appropriate for people with more severe depression who are at significant risk of suicide, harm to self or to others, self-neglect or complications in response to their treatment, leading to better outcomes and reduced need for more costly inpatient care.

The GC took into account the high costs associated with inpatient care, and decided to recommend inpatient treatment only for people with more severe depression who cannot be adequately supported by a crisis resolution and home treatment team.

Considering the benefits and costs of crisis resolution and home treatment teams (CRHTT) relative to other care settings, the GC expressed the opinion that CRHTT comprises an effective and likely cost-effective model of care for people with depression who would benefit from early discharge from hospital after a period of inpatient care.

The GC took into account the cost effectiveness of psychological treatments in the care of people with depression based on the results of the economic analysis undertaken for this guideline, and expressed the view that the full range of such treatments should also be available in inpatient settings, to allow provision of clinically and cost-effective care in

populations treated in such settings. The GC acknowledged the fact that increasing the intensity and duration of psychological interventions for people with depression in inpatient settings has resource implications, but expressed the view that the benefits of more intensive treatment in this group would outweigh the additional intervention costs. Moreover, if improved outcomes result in earlier discharge, then cost-savings may outweigh the intervention costs of more intensive psychological treatment.

The GC expressed the opinion that development of a treatment programme and a crisis management plan during contact with the CRHT team and on discharge or transfer to other services will allow more timely, appropriate and cost-effective planning and delivery of care to people with depression, that is targeted to their specific needs and thus can result in cost-savings (including a reduced rate of re-admission) that offset, fully or partially, any costs associated with the time spent on the development of the treatment programme. In contrast, lack of a detailed treatment programme and crisis management plan may lead to sub-optimal, less clinically and cost-effective care pathways and inappropriate treatments, ultimately leading to sub-optimal outcomes for the person and higher healthcare costs.

5.4.5.46 Quality of evidence

The GC noted that all outcomes had been assessed as either very low or low by GRADE. Most outcomes were downgraded due to indirectness, imprecision and risk of bias.

5.59 Recommendations

11. **Consider crisis and intensive home treatment for people with more severe depression who are at significant risk of:**
    - suicide, in particular for those who live alone
    - self-harm
    - harm to others
    - self-neglect
    - complications in response to their treatment, for example older people with medical comorbidities. [new 2017]

12. **Ensure teams providing crisis resolution and home treatment (CRHT) interventions to support people with depression:**
    - monitor and manage risk as a high-priority routine activity
    - establish and implement a treatment programme
    - ensure continuity of any treatment programme while the person is in contact with the CRHT team, and on discharge or transfer to other services when this is needed
    - have a crisis management plan in place before discharge from the team's care. [new 2017]

13. **Consider inpatient treatment for people with more severe depression who cannot be adequately supported by a CRHT team. [new 2017]**

14. **Make the full range of recommended psychological therapies (group CBT, CBT or BA) available for people with depression in inpatient settings. [new 2017]**

15. **When providing psychological therapies for people with depression in inpatient settings:**
    - increase the intensity and duration of the interventions

Update 2017

- ensure that they continue to be provided effectively and promptly on discharge. [new 2017]

16. Consider using CRHT teams with people with depression who might benefit from early discharge from hospital after a period of inpatient care. [2017]

6 Recognition and assessment

6.1 Introduction

The starting point for providing effective treatment for depression is the recognition of the problem and the first point of access is usually primary care, with the majority of people continuing to be managed in primary care. There is evidence, however, that many cases go unrecognised (Del Piccolo et al. 1998; Raine et al. 2000). Where depression is recognised, care often falls short of optimal recommended practice (Katon et al. 1992; Donoghue & Tylee 1996) and outcomes are correspondingly below what is possible (Rost et al. 1994). This is a cause of considerable concern. More recent studies, however, suggest that clinically significant depression (moderate to severe depressive illness) is detected by GPs at later consultations by virtue of the longitudinal patient–doctor relationship and it is milder forms, which are more likely to recover spontaneously, that go undetected and untreated (Thompson et al. 2001; Kessler et al. 2003).

In addition to efforts to improve recognition of depression, a number of responses have been developed over the past 20 or so years to address the problem of suboptimal treatment. These responses have included developments in the treatment of depression in primary and secondary care; the organisational and professional structures of primary and secondary care mental health services; and the development and adaptation of models for the management of chronic medical conditions, for example diabetes (Von Korff et al., 1997; Von Korff & Goldberg, 2001). Since the publication of the previous guideline in 2004, in the UK these developments have included the introduction of graduate mental health workers (Department of Health, 2003), which has contributed to increased access to low-intensity psychosocial interventions, including computerised CBT (NICE, 2006; NICE, 2005). The concept of ‘stepped care’ advocated in the previous guideline in 2004 has been embraced by many commissioners and providers in the NHS and is now being taken forward by the Improving Access to Psychological Therapies (IAPT) programme (Department of Health 2007; IAPT 2009). It is this later development, with £340 million of funding over 6 years along with 3,400 new psychological therapists that will bring about the single biggest change in the provision of effective treatments for depression in primary and secondary care. Since 2008 the IAPT programme in England has grown each year and in 2014/15 received more than 1.25 million referrals, and treated around 469,000 people, an estimated 15% of people with depression and anxiety disorders (HSCIC 2015).

This chapter focuses on one main issue: the identification of depression in primary and secondary care.

6.2 The identification of depression in primary care and community settings

6.2.1 Introduction

As stated above the accurate identification of depression is an essential first step in the management of people with depression. This includes both people who have sought treatment because of depressive symptoms and those being treated for other conditions, including physical health problems. The identification of depression in adults with a chronic physical health problem is covered in a related NICE guideline (NICE 2009). This guideline focuses on identifying depression in primary care and community settings.

Studies indicate that up to 50% of people with depression are not recognised when they attend primary care (Williams et al. 1995), a view which is supported by a recent meta-analysis of 37 studies of GPs’ unassisted ability to detect depression (Mitchell et al. 2009).

Mitchell and colleagues (2009) suggest that GPs are able to rule out depression in most people who are not depressed with reasonable accuracy but may have difficulty diagnosing depression in all true cases. However, as noted below, this under-recognition of depression may be focused more on mild depression than on moderate or severe depression (Kessler et al. 2003).

6.2.26 Identifying depression – a primary care perspective

For over 40 years, it has been suggested that GPs fail to accurately diagnose depression (Goldberg & Huxley 1992; Kessler et al. 2002). As stated above, some studies suggest that clinically important depression (moderate to severe depressive illness) is detected by GPs at later consultations by virtue of the longitudinal patient–doctor relationship and that its milder forms, which may recover spontaneously, go undetected and untreated (Thompson et al. 2000; Kessler et al. 2002). However, even this suggests that non-clinically important depression may go undetected initially. More recent studies suggest that the probability of prescribing antidepressants in primary care is associated with the severity of the depression, although almost half of the people prescribed antidepressants were not depressed (Kendrick et al. 2005). Other authors draw attention to the dangers of the erroneous diagnosis of depression in patients with a slight psychological malaise and few functional consequences that can lead to the risk of unnecessary and potentially dangerous medicalisation of distress (Aragones et al. 2006). Given the modest prevalence of depression in most primary care settings the number of false positive errors (people who are incorrectly identified as being at risk of depression) is larger than the number of false negatives (those falsely identified as not being at risk of developing depression). Further work is clearly needed to examine the subsequent outcome of those false positive and false negative diagnoses, and also to clarify the accuracy of GPs in diagnosing anxiety disorders, adjustment disorders and broadly defined distress.

Reasons for lack of recognition fall into four themes: factors related to the person with depression, and practitioner, organisational and societal factors.

6.2.28 Factors related to the person with depression

People may have difficulty in presenting their distress and discussing their concerns with their doctor, especially when they are uncertain that depression is a legitimate reason for seeing the doctor (Gask et al. 2003). The MaGPIe Research Group (2005a, 2005b) suggests that the relationship is important, and that GPs are, in fact, effective at identifying mental health problems in patients they know; however some people believe that the GP is not the right person to talk to, or that such symptoms should not be discussed at all. Negative perceptions about the value of consulting a GP for mental distress may, at least in part, explain low rates of help-seeking among young adults, including those with severe distress (Biddle et al. 2006). The person with depression may feel that they do not deserve to take up the doctor's time, or that it is not possible for doctors to listen to them and understand how they feel (Pollock & Grime 2002; Gask et al. 2003).

A number of other factors may also influence the identification of depression. Older adults, in particular, may complain less of depressed mood and instead somatise their depressive symptoms (Rabins 1996). Physical comorbidity can also make the interpretation of depressive symptoms difficult. People may have beliefs that prevent them from seeking help for depression such as a fear of stigmatisation, or that antidepressant medication is addictive or they may misattribute symptoms of depression for 'old age', ill health or grief. Although depression is more frequent in women, differential reporting of symptoms may lead to depression being under-diagnosed in men. From the perspective of the person with depression, it has been suggested that contact with primary care may be of little significance when set against the magnitude of their other problems (Rogers et al. 2001).

6.2.4 Practitioner factors

The construction of 'depression' as a clinical condition is contested amongst GPs (Chew-Graham et al. 2000, May et al. 2004, Pilgrim & Dowrick 2006). They may be wary of opening a 'Pandora's box' in time-limited consultations and instead collude with the person with depression in what has been called 'therapeutic nihilism' (Burroughs et al. 2006). In deprived areas, primary care physicians have been shown to view depression as a normal response to difficult circumstances, illnesses or life events (May et al. 2004), and depression may be under-diagnosed because of dissatisfaction with the types of treatment that can be offered, especially a lack of availability of psychological interventions. Primary care practitioners may also lack the necessary consultation skills or confidence to correctly diagnose late-life depression.

6.2.5 Organisational factors

The trend in the UK for mental health services to be separate from mainstream medical services may disadvantage people with depression who may have difficulties in attending different sites and/or services for mental and physical disorders.

Organisational factors that inhibit the identification and disclosure of symptoms and problems, together with limited access to mental health services, add to professionals' reluctance to encourage patients to disclose their distress (Popay et al. 2007, Chew-Graham et al. 2008).

6.2.6 Societal factors

The barriers described are likely to be particularly difficult for the economically poor and minority populations who tend to have more health problems and are more disabled. The oft-described barrier of stigma has to be set against the arguments that depression is a social construction within which chronic distress or unhappiness are medicalised (Ellis 1996, Pilgrim & Bentall 1999) and the suggestion that chronic unhappiness is not 'treatable' in the normal curative or therapeutic sense. It is therefore important that the healthcare professional recognises and accepts their own reaction to people presenting with depression so that they can acknowledge and go on to diagnose depression, and then discuss a range of possible interventions.

6.2.7 Shifting the emphasis from screening to identification

The identification of people with a disease is often referred to as screening (and was the term used in the previous 2004 guideline). Screening has been defined as the systematic application of a test or enquiry to identify individuals at high risk of developing a specific disorder who may benefit from further investigation or preventative action (Peckham & Dezateux 1998). Screening programmes detect people at risk of having the condition or at risk of developing the condition in the future. They do not establish a diagnosis but give some indication of any action that may be required, such as further diagnostic investigation, closer monitoring or even preventative action. Screening is not necessarily a benign process (Marteau 1989). Since screening tools are never 100% accurate, people who are incorrectly identified as being at risk of developing a condition (false positives) can be subject to further possibly intrusive, harmful or inappropriate investigations, management or treatment. Those falsely identified as not being at risk of developing a condition (false negatives) will also suffer by not being given the further investigation they need.

Critics of routine screening for depression have advanced a number of arguments against it. These include the low positive predictive value of the instruments (that is, many patients who screen positive do not have depression), the lack of empirical evidence for benefit to patients, the expenditure of resources on patients who may gain little benefit (many patients who are detected by such an approach may be mildly depressed and recover with no formal

intervention), and the diversion of resource away from more seriously depressed and known patients who may be inadequately treated as a result. These issues are well covered by Palmer and Coyne (2003) in their review of screening for depression in medical settings. Palmer and Coyne (2003) also go on to make a number of suggestions for improving recognition, including ensuring effective interventions for those identified, focusing on patients with previous histories of depression and people known to have a high risk of developing depression, such as those with a family history of the condition or chronic physical health problems with associated functional impairment. Others (for example, Pignone et al. 2002, Macmillan et al. 2005) have, however, recommended the use of screening of depression for the general adult population, but it should be noted that the systematic review of interventions conducted in support of the recommendations by these groups have included the need for follow-up interventions. The effectiveness of such interventions (for example, feedback to patients or case management) is considered below and the GDG felt it important to first address the value of case identification systems alone, before going on to consider the benefits of integrated systems.

Within the NHS, between 2006 and 2013, case identification of depression in people with diabetes and ischaemic heart disease was part of routine clinical work for primary care practitioners as stipulated by the GP Contract Quality and Outcomes Framework (BMA & NHS Employers 2006), using the two-item Whooley questions, which have high sensitivity in the detection of depression (Bosanquet et al. 2015). It has been suggested that using an additional question ('is this something with which you would like help?' [Arroll et al. 2005]) may improve the specificity of the screening questions, but the current evidence for the use of an additional help question is not consistent and there is, as yet, insufficient data to recommend its use for screening or case finding (Bosanquet et al., 2015).

Update 2017

Others, however, caution that the use of such screening instruments may encourage practitioners to take a reductionist, biomedical approach, diverting them from a broader bio-psychosocial approach to both diagnosing and managing depression (Dowrick 2004).

6.3 Case identification

6.3.1 Introduction

The previous NICE guideline on depression, in addition to other NICE mental health guidelines, considered the case for general population screening for a number of mental health disorders and concluded that it should only be undertaken for specific high-risk populations where benefits outweigh the risks (for example, NICE 2011). These were people with a history of depression, significant physical illnesses causing disability, or other mental health problems, such as dementia.

A history of depression has been identified as a significant factor in future episodes. For example, a study of 425 primary care patients found that 85% of those who were depressed had had at least one previous episode (Coyne et al.1999). In fact, having a history of depression produced a positive predictive value (see below) roughly equal to that produced by using a depression case-finding instrument (Centre of Epidemiology Studies-Depression – CES-D) (0.25 compared with 0.28). This suggests that careful assessment of relevant instruments is required if a number currently in use appears to have no more predictive value than a history of depression. It should be noted that depression can frequently be comorbid with other mental health problems, including borderline personality disorder (for example, Zanarini et al.1998, Skodol et al.1999), and dementia (Ballard et al.1996).

The following sections review available case identification instruments.

6.3.21 Definition

Case identification instruments were defined in the review as validated psychometric measures that were used to identify people with depression. The review was limited to identification tools likely to be used in UK clinical practice, that is, the Beck Depression Inventory (BDI), Patient Health Questionnaire (PHQ), General Health Questionnaire (GHQ), Centre of Epidemiology Studies-Depression (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), Zung Self Rated Depression Scale and any one- or two-item measures. The identification tools were assessed in consultation (which included primary care and general medical services) and community populations. 'Gold standard' diagnoses were defined as DSM–IV or ICD–10 diagnosis of depression. Studies were sought that compared case identification with one of the above instruments with diagnosis of depression based on DSM–IV or ICD–10 criteria. Studies that did not clearly state the comparator to be DSM–IV or ICD–10, used a scale with greater than 28 items, or did not provide sufficient data to be extracted in the meta-analysis were excluded.

6.3.35 Summary statistics used to evaluate identification instruments

Sensitivity, specificity, positive predictive validity and negative predictive validity

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 negatives it 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), 1,000 people are given a test which has 90% sensitivity and 85% specificity. It is known that 100 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).

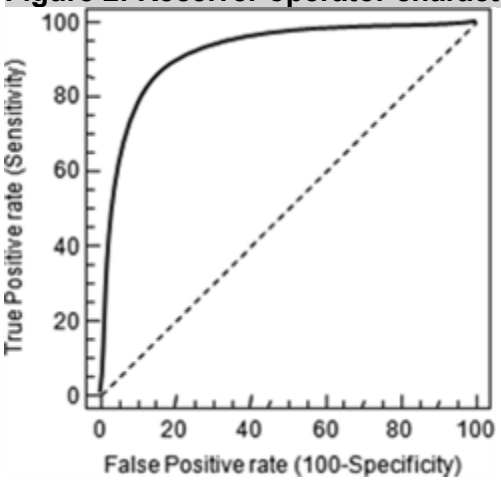
On the other hand, sensitivity and specificity do not necessarily depend on prevalence of depression (Altman & Bland 1994b). For example, sensitivity is concerned with the performance of an identification test conditional on a person having depression. Therefore the higher false positives often associated 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 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 operator characteristic curves

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

Figure 2: Receiver operator characteristic curve



A test with perfect discrimination would have an 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.

Negative and positive likelihood ratios

Negative (LR-) and positive (LR+) likelihood ratios are thought not to be dependent on prevalence. LR- is calculated by sensitivity/1-specificity and LR+ is 1-sensitivity/ specificity. A value of LR+ >5 and LR- <0.3 suggests the test is relatively accurate (Fischer et al. 2003).

Diagnostic odds ratios

The diagnostic odds ratio is LR+/LR-; a value of 20 or greater suggests a good level of accuracy (Fischer et al. 2003).

6.3.41 Databases searched and inclusion/exclusion criteria

The review team conducted a new systematic search for cross-sectional studies to assess tools for identifying depression. This was undertaken as a joint review for this guideline and the guideline for depression in adults with a chronic physical health problem (NICE 2009c). Information about the databases searched and the inclusion/exclusion criteria used can be found in Table 37. Details of the search strings used are in Appendix H.

Table 37: Databases searched and inclusion/exclusion criteria for the effectiveness of case identification instruments

Electronic databases	MEDLINE, EMBASE, PsycINFO, Cochrane Library
Date searched	Database inception to February 2009
Study design	Cross-sectional studies
Patient population	People in primary care, community, and general hospital settings
Instruments	BDI, PHQ, GHQ, CES-D, GDS, HADS, Zung Self Rated Depression Scale, and any one- or two-item measures of depression
Outcomes	Sensitivity, specificity, AUC, diagnostic odds ratio, positive likelihood, negative likelihood

6.3.59 Studies considered

A total of 126 studies met the eligibility criteria of the review; 54 studies were conducted in consultation samples, 45 were on people with chronic physical health problems[b] and 50 were on older people (over 65 years of age). Of these studies, 16 were on the PHQ-9, five on the PHQ-2, six on the 'Whooley questions', 19 on the BDI, nine on the BDI – short form, two on the GHQ-28, 12 on the GHQ-12, 17 on the CES-D, 20 on the GDS, 11 on the GDS-15, 16 on HADS-D, five on HADS-total and seven on one-item measures (see Appendix J2 for further details).

In addition, 251 studies were excluded from the analysis. The most common reason for exclusion was a lack of a gold standard (DSM/ICD) comparator (see Appendix J2 for further details).

6.3.60 Evaluating identification tools for depression

A bivariate diagnostic accuracy meta-analysis was conducted using Stata 10 with the Module for Meta-analytical Integration of Diagnostic Test Accuracy Studies (MIDAS) (Dwamena 2007) commands in order to obtain pooled estimates of sensitivity, specificity, likelihood ratios and diagnostic odds ratio. To maximise the available data, the most consistently reported and recommended cut-off points for each of the scales were extracted (see Table 38).

Table 38: Cut off points used (if available) for each of the identification tools (adapted from Pignone et al. 2002; Gilbody et al. 2007)

Scale	Cut off points
BDI 21 items 13 items Primary care version	13 4 4
PHQ 9 items	10 3

b Data for the population with chronic physical health problems and information about the included studies is presented in the related guideline, Depression in Adults with a Chronic Physical Health Problem (NICE 2009).

Scale	Cut off points
2 items 2 items (Whooley version)	1
GHQ 28 items 12 items	5 3
HADS-D	8–10 mild, 11–14 moderate, 15+ severe
CES-D	16
GDS 30 items 15 items 5 items	10 5 ?
Zung	50 mild, 60 moderate, 70 severe

Heterogeneity is usually much greater in meta-analyses of diagnostic accuracy studies compared with RCTs (Gilbody et al. 2007; Cochrane Collaboration 2008). Therefore, a higher threshold for acceptable heterogeneity in such meta-analyses is required. However when pooling studies resulted in $I^2 > 90\%$, meta-analyses were not conducted.

Table 39 summarises the results of the meta-analysis in terms of pooled sensitivity, specificity, positive likelihood ratios, negative likelihood ratios, and diagnostic odds ratios. Additional subgroup analyses were conducted for older adults.

Table 39: Evidence summary of depression identification instruments in primary care, people with a chronic physical health problem, and older populations

Population and instrument	Sensitivity	Specificity	Likelihood ratio+	Likelihood ratio	Diagnostic odds ratio	AUC
PHQ-9 Consultation samples: 11 studies	0.82 (0.77, 0.86)	0.83 (0.76, 0.88)	4.70 (3.29, 6.72)	0.22 (0.17, 0.29)	21.38 (11.87, 38.52)	0.88 (0.85, 0.91)
Whooley*: All populations: 7 studies	0.95 (0.91, 0.97	0.66 (0.55, 0.76)	2.82 (2.01, 3.96)	0.08 (0.04, 0.15)	36.25 (14.89, 88.24)	0.94 (0.92, 0.96)
BDI Consultation samples: 4 studies	0.85 (0.79, 0.90)	0.83 (0.70, 0.91)	5.14 (2.83, 9.32)	0.18 (0.12, 0.24)	29.29 (15.10, 56.79)	0.90 (0.87, 0.92)
BDI-non somatic items Consultation sample: 5 studies	0.82 (0.57, 0.94)	0.73 (0.61, 0.83)	3.02 (1.87, 4.90)	0.25 (0.09, 0.69)	11.92 (3.02, 47.04)	0.83 (0.79, 0.86)
CES-D Consultation sample: 8 studies Older adults: 5 studies	0.84 (0.78, 0.89) 0.81 (0.74, 0.87)	0.74 (0.65, 0.81) 0.79 (0.67, 0.88)	3.19 (2.41, 4.22) 3.82 (2.35, 6.22)	0.21 (0.15, 0.29) 0.24 (0.17, 0.33)	15.02 (9.38, 24.05) 15.95 (8.05, 31.60)	0.87 (0.84, 0.90) 0.83 (0.80, 0.86)
GDS-15 Consultation sample: 11 studies	0.87 (0.80, 0.91)	0.75 (0.69, 0.80)	3.40 (2.73, 4.24)	0.18 (0.12, 0.27)	18.98 (10.85, 33.20)	0.86 (0.83, 0.89)
1-item Consultation sample: 6 studies	0.84 (0.78, 0.89)	0.65 (0.55, 0.73)	2.38 (1.81, 3.13)	0.25 (0.17, 0.36)	9.67 (5.35, 17.46)	(0.82, 0.88)
Notes: *It was not possible to conduct separate subgroup analyses for consultation and chronic physical illness samples due to lack of studies for the Zung and Whooley questions.						

Patient Health Questionnaire

The PHQ developed out of the more detailed Primary Care Evaluation of Mental Disorders (PRIME-MD) (Spitzer et al.1994). There are three main instruments that have been developed from this scale; the PHQ-9 (Spitzer et al.1999), PHQ-2 (Kroenke et al. 2003) and the 'Whooley questions' (Whooley et al.1997).

The PHQ-9 has nine items and has a cut-off of 10. Although the PHQ-2 and the Whooley questions use the same two items, the difference is that while the PHQ-2 follows the scoring format of the PHQ-9 (Likert scales), the Whooley version dichotomises the questions (yes/no) and has a cut-off of 1 compared with 3 for the PHQ-2.

For the PHQ-9 in consultation samples (people in primary care or general medical settings) there was relatively high heterogeneity (although of a similar level to most other scales) (I^2 = 74.04%). The PHQ-9 was found to have good sensitivity (0.82, 95% CI, 0.77, 0.86) and specificity (0.83, 95% CI, 0.76, 0.88).

The PHQ-2 could not be meta-analysed as there was very high heterogeneity. The Whooley questions analysis included studies both on consultation and chronic physically ill samples as there were too few studies to break down by population. This scale was found to have high sensitivity (0.95, 95% CI, 0.91, 0.97) but lower specificity (0.66, 95% CI, 0.55, 0.76). A single study by Arroll and colleagues (2005) added a further question to the two in the PHQ-2, asking the patient if they wanted help with their depression. This increased specificity and the GDG considered the findings of the study and the adoption of the third question, but as there was only a single study showing the effect of this approach the GDG decided not to adopt it.

It was not possible to conduct meta-analysis on the effects of any of the PHQ scales or the Whooley questions on older adults because of a lack of data (one study each on the PHQ-9, PHQ-2 and Whooley questions).

Beck Depression Inventory

Beck originally developed the BDI in the 1960s (Beck et al.1961) and subsequently updated the original 21-item version (Beck et al., 1979; Beck et al. 1996). This scale has been used widely as a depression outcome measure and is also used to provide data on the severity of depression; commonly, 13 is used a cut-off in identification studies. In addition, the cognitive–affective subscale of the BDI has often been used to identify depression. Furthermore, the BDI-fast screen has been specifically developed for use in primary care (Beck et al. 1997).

For the 21-item BDI there was high heterogeneity for consultation samples (I^2 = 88.61%). The BDI appeared to perform relatively well in terms of sensitivity (0.85, 95% CI, 0.79, 0.90) and specificity (0.83, 95% CI, 0.70, 0.91). This was also consistent with the diagnostic odds ratio (29.29, 95% CI, 15.103, 56.79). However, this is based on only four studies so it is difficult to draw firm conclusions. Subgroup analyses on older adults were also not possible as there were only two studies for this population.

Beck Depression Inventory – non-somatic items

Data from BDI fast-screen (Beck et al. 2000) and BDI short-form (Beck et al. 1974, 1996) were combined to assess the impact of removing somatic items as data from both scales were relatively sparse. There was sufficient, although relatively low, consistency between studies to assess these scales (BDI: non-somatic) in consultation (I^2 = 75.71%) populations. There was high sensitivity (0.82, 95% CI, 0.57, 0.94) but lower specificity (0.73, 95% CI, 0.61, 0.83). A meta-analysis was not possible for older adults as there were only two studies.

General Health Questionnaire

The GHQ (Goldberg & Williams 1991) was developed as a general measure of psychiatric distress and measures a variety of constructs such as depression and anxiety. The main versions used for identification purposes are the GHQ-28 (cut-off of 5) and GHQ-12 (cut-off of 3).

There were only two trials of the GHQ-28, therefore meta-analysis was not conducted. In addition, while there were more studies on the GHQ-12 there was very high heterogeneity (I^2 > 90%) for studies on consultation populations, therefore these studies were also not meta-analysed. Moreover, a meta-analysis specifically for older adults was not possible due to there being only two studies.

Hospital Anxiety and Depression Scale

The HADS (Zigmond & Snaith 1983) is a measure of depression and anxiety developed for people with physical health problems. The depression subscale has seven items and the cut-off is 8 to 10 points.

A total of 21 studies were included in the review, however meta-analysis could not be conducted due to very high heterogeneity (I^2 > 90%) for all subgroups including consultation populations and older adults.

Center for Epidemiological Studies Depression Scale

The CES-D (Radloff 1977) has 20 items and the cut-off is 16. This measure is also relatively commonly used as an outcome measure. There are various short forms of the CES-D including an eight-, ten- and 11-item scale.

There was high heterogeneity in the consultation (I^2 = 84.63%) sample. For the older adult population, Haringsma and colleagues (2004) was removed from the analysis resulting in acceptable heterogeneity (I^2 = 61.09%).

For consultation samples sensitivity was high (0.84, 95% CI, 0.78, 0.89) but specificity was lower (0.74, 95% CI, 0.65, 0.81). For older adults, there was relatively low sensitivity (0.81, 95% CI, 0.74, 0.87) and higher specificity (0.79, 95% CI, 0.67, 0.87).

Geriatric Depression Scale

The GDS was developed to assess depression in older people. The original 30-item scale (cut-off of 10 points) was developed by Yesavage and colleagues (1982) and more recently a 15-item (cut-off of 5 points) version has been validated.

Despite the large number of studies (18 studies), there was very high heterogeneity (I^2 > 90%) for the GDS, therefore no meta-analyses could be conducted. However, it was possible to analyse studies on the GDS-15.

In the consultation population there was higher sensitivity (0.87, 95% CI, 0.80, 0.91) but specificity was relatively low (0.75, 95% CI, 0.69, 0.80). The diagnostic odds ratio was just below 20 (18.98, 95% CI, 10.85, 33.20). Heterogeneity was relatively acceptable (I 2 = 70.96%).

No subgroup analyses for older people were conducted as all participants were over 65 years of age.

Zung Self-Rating Depression Scale

The self-rating depression scale was developed by Zung (Zung, 1965) and has been revised (Guy, 1976). This has 20 items where a cut-off of 50 is typically used. It is sometimes used as an outcome measure as well. There were insufficient studies to conduct a meta-analysis.

One-item measures

Five studies were found to assess a one-item measure in consultation samples. There was a relatively good sensitivity (0.84, 95% CI, 0.78, 0.89) but very low specificity (0.65, 95% CI, 0.55, 0.73). The diagnostic odds ratio indicated a lack of accuracy.

(9.67, 95% CI, 5.35, 17.46). It was not possible to conduct a subgroup analysis of older adults as there were only two studies.

Comparing validity coefficients for case identification tools in older adults

The impact of old age and residing in a nursing home on the validity coefficients of the case identification tools reviewed above were assessed through meta-regression (see Table 40). Because of a lack of data the PHQ-2, Whooley, Zung, and one-item measures were not included in the analysis.

The GDS and GDS-15 were almost always used for older adults, therefore the validity of these measures in older adults is already accounted for in the previous analysis. However, further analyses were conducted to assess the validity of these measures in nursing home populations.

Table 40: Meta-regressions assessing the impact of differences within populations of studies

Population and instrument	Beta-coefficient	I^2(%)	p-value
PHQ-9 Comparing over 65s with under 65s	Sensitivity = 1.23 Specificity = 1.84	Joint I^2 = 0	0.65 0.73 0.83
BDI Comparing over 65s with under 65s	Sensitivity = 1.58 Specificity = 0.74	Joint I^2 = 0	0.34 0.79 0.65
BDI-non somatic items Comparing over 65s with under 65s	Sensitivity = 1.58 Specificity = 2.12	Joint I^2 = 58.64	0.80 0.02 0.09
CES-D Comparing over 65s with under 65s	Sensitivity = 1.23 Specificity = 1.61	Joint I^2 = 43.30	0.09 0.18 0.17
GDS Comparing nursing home with non-nursing home	Sensitivity = 1.54 Specificity = 1.13	Joint I^2 = 0	0.85 0.65 0.80
GDS-15 Comparing nursing home with non-nursing home	Sensitivity = 2.14 Specificity = 0.91	Joint I^2 = 0	0.36 0.34 0.44
GHQ-12 Comparing over 65s with under 65s	Sensitivity = 0.43 Specificity = 1.45	Joint I^2 = 11.28	0.14 0.33 0.32

Older adults

There was some evidence that the BDI versions with no somatic items (p = 0.02) were associated with improved specificity in older adults compared with people under 65 years. There was a trend towards reduction in sensitivity for the CES-D (p = 0.09) in older adults compared with people under 65 years. For all other scales there were no statistically significant differences. However, there was often a lack of power in most studies because only a small number of studies on older adults were found for most scales.

People in nursing homes

Only the GDS and GDS-15 provided sufficient data on people in nursing homes. There appeared to be limited differences in validity when assessing people either in nursing homes or in the community for both scales.

6.4 Case identification in black and minority ethnic populations

6.4.1 Introduction

Culture and ethnicity are known to influence both the prevalence and incidence of mental illnesses, including common mental disorders such as depression (Bhui et al. 2001). For example, Shaw and colleagues (1999) indicated that women from black and minority ethnic groups had an increased incidence of common mental disorders including both depression and anxiety. Such findings cannot wholly be explained by differences in factors such as urbanicity, socioeconomic status and perceptions of disadvantage (Bhugra & Cochrane 2001, Weich et al. 2004). Furthermore, culture is known to exert an influence on the presentation and subjective experience of illness. What a person perceives as an illness and whom they seek for treatment are all affected by their culture and ethnicity. With regard to depression, a number of findings have indicated both ethnic and cultural variations in the subjective experience and initial presentation of the illness. For example, Commander and colleagues (1997) are among researchers who suggest that 'Asians', including Indian, Bangladeshi and Pakistani people, are more likely to present to their GP with physical manifestations, and do so more frequently than their white counterparts. However, both Wilson and MacCarthy (1994) and Williams and Hunt (1997) have indicated that despite this increased GP contact, and even when a psychological problem is present, GPs are less likely to detect depression and more likely to diagnose 'Asians' with a physical disorder.

There is an increasing evidence base to suggest that the reduced identification of depression in different ethnic and cultural groups may be one barrier to receiving appropriate treatment, including both psychological and pharmacological interventions. For example, research has suggested that across mental disorders, particular ethnic groups are often under-represented in primary care services (Bhui et al. 2003; Department of Health 2008b), whereas a Healthcare Commission survey highlighted how both Asian and black/black British people were less likely to be offered 'talking therapies' (Department of Health 2008b).

Despite an increased awareness that different cultural and ethnic factors may influence the presentation of depression, the majority of case identification tools used in routine clinical practice were originally created and validated in white populations (Husain et al. 2007). Owing to the above evidence indicating ethnic and cultural variations in the presentation and subjective experience of illness, one proposed method to improve the identification of depression in black and minority ethnic participants is to assess the validity of ethnic-specific screening tools. Such tools, most of which are still early in their development, aim to incorporate specific cultural idioms and descriptions commonly reported by people from a particular ethnic or cultural group.

6.4.21 Definition and aim of topic review

The review considered any ethnic-specific case identification instruments aimed at detecting depression in black and minority ethnic populations. This included new identification tools designed for different cultural and ethnic groups, and also existing scales modified and tailored towards the specific needs of particular black and minority ethnic groups. Although the GDG was aware of papers from outside the UK (most notably from the US), the decision was made to only include UK studies. As discussed above, the presentation and subjective experience of depression is known to be influenced by cultural and ethnic factors; therefore, it was felt that findings from non-UK ethnic minority populations would not be generalisable because of the ethnic and cultural differences among the populations studied. The review also assessed the validity of established depression case identification tools for different black and minority ethnic populations within the UK[c].

6.4.33 Databases searched and inclusion/exclusion criteria

The review team conducted a new systematic search for cross-sectional studies aiming to assess tools for identifying depression. This was undertaken as a joint review for this guideline and the guideline for depression in adults with a chronic physical health problem (NCCMH, 2010). Information about the databases searched and the inclusion/exclusion criteria used are presented in Table 41.

Table 41: Databases searched and inclusion/exclusion criteria for clinical effectiveness of psychological interventions

Electronic databases	MEDLINE, EMBASE, PsycINFO, Cochrane Library
Date searched	Database inception to February 2009
Study design	Cross-sectional studies
Patient population	People in primary care, community, and general hospital settings from black and minority ethnic groups
Instruments	1. Any ethnic-specific depression case identification instrument 2. Any cultural or ethnically adapted version of the following validated case identification instruments: BDI, PHQ, GHQ, CES-D, GDS, HADS, Zung Self Rated Depression Scale, and any one- or two-item measures of depression 3. Any of the above validated identification tools, assessed in a UK black and minority ethnic population
Outcomes	Sensitivity, specificity, AUC, diagnostic odds ratio, positive likelihood, negative likelihood

6.4.421 Studies considered

A total of four studies met the eligibility criteria of the review. All four papers were conducted within the community or primary care. One included study compared the Amritsar Depression Inventory (ADI) with the GHQ-12, and two studies compared the Caribbean Culture-Specific Screen for emotional disorders (CCSS) with the GDS. Only one study assessed the validity of an established scale, the Personal Health Questionnaire, in a UK black and minority ethnic population, namely people of Pakistani family origin.

In addition, ten studies were excluded from the analysis. The most common reason for exclusion was that the paper was a non-UK based study/population or that the paper presented no usable evaluation of a screening tool.

c Papers assessing the validity of established scales in UK black and minority ethnic populations were required to have a 'gold standard' diagnosis defined as DSM–IV or ICD–10 diagnosis of depression

6.4.51 Evaluating identification tools for depression in black and minority ethnic 2 populations

3 Because of both the paucity of data on ethnic specific scales in the UK and differences in the
4 populations and instruments investigated, it was not possible to conduct a meta-analysis of
5 the included studies. Instead the findings from the included studies are summarised in a
6 narrative review below.

7 Amritsar Depression Inventory

8 The ADI is a culturally specific instrument developed in the Punjab in India and is aimed at
9 detecting depression in the Punjabi population of the Indian subcontinent (Singh et al., 1974).
10 The 30-item dichotomous (yes/no) questionnaire was developed on the basis of 50
11 statements commonly used by Punjabi people with depression. The screen development
12 process also utilised frequently used 'illness statements' and common descriptions of signs
13 and symptoms of depression prevalent in the psychiatric literature.

14 Using the ADI and the GHQ-12, Bhui and colleagues (2000) screened both Punjabi and
15 white English attendees of five primary care practices in South London. Throughout the
16 study, a cultural screen assessing self-affirmed cultural origin was applied to detect both
17 Punjabi and white English participants. To overcome any additional barriers because of
18 language, the screening tools were administered in English, Punjabi or a combination of the
19 two, depending on the preference of the participant. A two-phase screening protocol was
20 applied in which all 'probable cases', for example, those scoring >2 on the GHQ or >5 on the
21 ADI, and one third of 'probable non-cases' proceeded to a second interview in which the
22 Clinical Interview Schedule-Revised (CIS-R) was administered by a bilingual psychiatrist.

23 Results of the validity coefficients and ROC curve analysis using the standard CIS-R
24 thresholds for depression indicated that while the GHQ-12 performed well across both
25 groups, culture had an impact on the validity coefficient of the ADI. In particular, although
26 performing in line with the GHQ-12 for the white English participants, the ADI performed
27 worse in detecting depression in the Punjabi participants. Results indicated that the ADI was
28 no better than chance in identifying cases of depression, particularly for Punjabis who had
29 been resident in the UK for more than 30 years. One additional finding of interest was that
30 the optimal cut-off for the ADI was higher for the Punjabi participants compared with their
31 white English counterparts, although this finding was not sustained for the GHQ-12 in which
32 the same cut-off was optimal for both groups. Analysis of the individual items of both the
33 GHQ-12 and the ADI failed to indicate any specific items that were strongly predictive of
34 depression caseness in either cultural group.

35 Caribbean Culture-Specific Screen for emotional distress

36 The CCSS (Abas 1996) is a 13-item dichotomous (yes/no) culture-specific screen which was
37 developed through a process of generating locally-derived classifications of mental disorders
38 in Caribbean people and gathering commonly used terms for emotional distress. The
39 majority of participants interviewed in the piloting stages of the screen were from Jamaica
40 with a number of participants identifying themselves as from other Caribbean countries
41 including Guyana, Barbados, Trinidad and Grenada.

42 Two papers assessed the validity of the CCSS screen in older African–Caribbean
43 participants living in two different locations in the UK, namely South London and Manchester.
44 Both papers compared the validity of the CCSS to the GDS and utilised the Geriatric Mental
45 State-Automated Geriatric Examination for Computer Assisted Taxonomy (GMS-AGECAT)
46 as a gold standard for case identification.

47 The sample in Abas and colleagues (1998) consisted of consecutive African-Caribbean
48 primary care users aged over 60, and included both clinic attendees and those receiving

home visits from primary care teams. Participants were firstly administered the CCSS, GDS-15 and the Mini-Mental State Examination (MMSE). Responders were categorised as high scorers if they scored >4 on either measure, and low scorers if they attained less than 4 on both screens. A random sample of 80% of the high scorers and 20% of the low scorers was selected to attend a further interview. During this second stage interview, the GMS-AGECAT and a culturally-specific diagnostic interview, which was informed through a process of consultation with African–Caribbean religious healers/ministers, were administered to the selected participants.

Rait and colleagues (1999) included a community sample of African–Caribbean people aged 60 years and over. Registers for general practices with a high-proportion of African–Caribbeans were used to identify members of the community. In stage one, letters were sent to potential participants, with those who consented to take part in the study subsequently interviewed in their homes. All included participants were interviewed by one of two interviewers of a similar cultural background. During this stage, three depression screens were applied, namely the GDS-15, CCSS and the Brief Assessment Schedule Depression Cards (BASDEC). The second stage of the study involved the home administration of the GMS-AGECAT, used as a diagnostic 'gold standard' for the detection of depression.

The ROC curve analyses for the papers indicated that both the GDS and the CCSS performed well in the populations, with a high level of sensitivity and specificity when using the GMS-AGECAT as a gold standard for diagnosis. In both papers, the culturally-specific CCSS did not outperform the GDS. In the Abas and colleagues' (1998) paper it was demonstrated that at a certain cut-off the GDS appeared to perform better than the CCSS, although the authors noted that the small sample size prevented any meaningful test of statistical significance. Because it was noted that considerable variation may exist among people of Caribbean origin from different islands, for example, Jamaica, Trinidad and so on, the results of Rait and colleagues' (1999) paper were presented for the sample as a whole and for a subgroup of Jamaican people who constituted the majority of participants. Although slight variation existed between the two analyses, the results were similar, with the same optimal cut-off occurring in both analyses.

One important feature of the Rait and colleagues' (1999) study was that the authors sought advice from a panel of community resident African–Caribbeans regarding the acceptability of the GDS. The content of the screens was deemed acceptable, and no suggestions for changes were made. Rait and colleagues (1999) argue that the success of case identification measures may be more dependent on the way in which the screen is delivered, for example, the cultural competence of staff and delivering the screen in a culturally sensitive way, rather than the content per se. This conclusion was supported by Abas and colleagues (1998) who found that a proportion of participants were more likely to discuss and disclose information during the culturally sensitive diagnostic interview, when compared with the standard GMS-AGECAT. Consequently, both papers have suggested that routine clinical screens may be appropriate for black and minority ethnic participants, particularly when delivered in a culturally sensitive way.

Personal Health Questionnaire

Husain and colleagues (2007) assessed the validity of the Personal Health Questionnaire in Pakistani people who were resident in the UK. The authors noted that, unlike many screening instruments, the Personal Health Questionnaire contains no 'difficult culture specific idioms', thus making translations into other languages possible. In the present study, the Personal Health Questionnaire was translated and back-translated into Urdu, the main language of immigrants from Pakistan, with group discussion utilised to reach a single consensus.

Consecutive primary care attendees of Pakistani family origin aged 16 to 64 years were included in the sample. Eligible participants were identified through either their name and/or

language or via direct questioning. As with the other screening studies, a two stage process was employed. All eligible participants first completed the Personal Health Questionnaire in either English or Urdu, depending on patient preference, with a research psychiatrist administering the screen in the case of illiteracy. In the second stage of the study, all participants were interviewed in either their home or within the primary care practice. A psychiatrist administered the Psychiatric Assessment Schedule, a semi-structured interview resulting in an ICD diagnosis, in either Urdu or English dependent on preference.

Results of the ROC curve analysis indicated that the recommended cut off score of >7 produced a sensitivity of 70.4% and a specificity of 89.3%, with a positive predictive value of 82.6 and a negative predictive value of 80.6. The high sensitivity and specificity at the recommended cut-off suggested that the Personal Health Questionnaire is able to detect depression in people of Pakistani family origin when administered in either English or Urdu. Furthermore, the authors noted that participants in this study and in a study conducted in Pakistan (Husain et al. 2000) did not experience any difficulties in understanding and answering the screening questions.

6.4.6 Limitations with the evidence base

It must be noted that a number of potential limitations exist in relation to the above studies. One caveat is the lack of an established gold standard for the diagnosis of depression in people from black and minority ethnic groups. Only one paper used a culturally-sensitive diagnostic tool as a measure of caseness (Abas et al. 1998). The remaining three papers compared the screens with long-standing measures predominantly based on the DSM and ICD–10 classification systems. It is argued that these measures may not be culturally specific and sensitive to cultural differences, but are instead based on ethnocentric ideas of mental illness (Bhui et al. 2000). Consequently, any culturally sensitive measure may not be expected to have a high sensitivity and specificity for caseness when compared with these diagnostic measures. Further research into this area is therefore required to answer such questions.

A further caveat to consider is that three of the four studies that were included assessed consecutive primary care attendees, who may or may not be wholly representative of ethnic minorities, particularly those who experience barriers to accessing and engaging with primary care services. However, the findings of one paper in which a community sample was recruited were consistent with the results of the primary care studies, suggesting the findings may be robust for each particular ethnic group under investigation.

6.5 Clinical summary for both reviews

There was very high heterogeneity found for almost all identification tools, which is an important limitation of the reviews. Scales varied a great deal in terms of targeted populations, number of items and scoring systems. When compared with the Whooley questions, other scales such as the PHQ-9 and GDS-15 had better specificity but not as much sensitivity (although they still met the criteria for high sensitivity).

There were also planned subgroup analyses conducted for older adults, which included scales specifically targeted at this population (for example, the GDS and GDS-15) as well as all other measures reviewed. The GDS-15 appeared to be relatively effective in consultation populations. However, the large number of studies on the 30-item GDS could not be meta-analysed as there was very high heterogeneity. There were fewer studies on the CES-D, but the available data suggested a slightly (although not statistically significant) reduced sensitivity compared with consultation populations as a whole. There were studies that targeted older adults for all of the other scales reviewed; however, the number of studies was too small to conduct meta-analyses for any of these measures.

There was a paucity of data concerning ethnic-specific identification tools, with limited data suggesting that the scales, which may be in their developmental infancy, failed to detect depression in different ethnic and cultural groups. In all studies, validated and well researched measures such as the GHQ-12 outperformed the ethnic-specific scales in terms of both sensitivity and specificity. Furthermore, in the case of the Personal Health Questionnaire, this was validated in a particular black and minority ethnic group, namely Pakistani people resident in the UK.

6.6 Health economic evidence and considerations

No evidence on the cost effectiveness of case identification tools for depression in primary care and community settings was identified by the systematic search of the economic literature.

6.7 From evidence to recommendations

The GDG noted the different nature of the scales contained in the review and their psychometric properties, as well as the possible benefit of a two-stage process of identification and diagnosis.

The first stage of case identification would require using a highly sensitive instrument that could be used in routine clinical practice with limited training and implementation difficulties. The data supported the use of the Whooley questions and, given that this measure is already in current use in primary care, the GDG concluded that in the first stage of case identification the Whooley questions remained an appropriate tool for depression. However, given the lack of specificity found with the Whooley questions it was the view of the GDG that people with a positive response would benefit from a more detailed clinical assessment, which may include a more detailed instrument possessing better overall psychometric properties. The data on case-finding instruments in black and minority ethnic groups did not identify any specific measures that in the opinion of the GDG improved upon the results obtained with the Whooley questions, and therefore no specific black and minority ethnic recommendations on case finding tools are made. However, the need for cultural competence of staff in assessments was noted in the review of case-finding instruments in black and minority ethnic groups, and this is reflected in the recommendations. In addition, in performing a more comprehensive mental health assessment, as recommended in the previous 2004 guideline, the need to move beyond simple symptom counts was noted, so the recommendation from the previous 2004 guideline has been amended. This guideline update also makes recommendations for people with depression and learning disabilities or acquired cognitive impairments because it is likely that depression, which is ‘relatively common’ (Prasher 1999) in this population, will be under-diagnosed, particularly if they have autism, a learning disability, established aggressive, self-harming or over-active behaviours or comorbid physical health problems such as epilepsy, diabetes or heart disease (Prasher, Mind 2007). Other recommendations from the previous 2004 guideline remain essentially the same.

6.8 Recommendations

17. Be alert to possible depression (particularly in people with a past history of depression or a chronic physical health problem with associated functional impairment) and consider asking people who may have depression if:

- During the last month, have they often been bothered by feeling down, depressed or hopeless?
- During the last month, have they often been bothered by having little interest or pleasure in doing things? [2009]

18. If a person answers 'yes' to either of the depression identification questions (see recommendation 22) but the practitioner is not competent to perform a mental health assessment, refer the person to an appropriate professional who can. If this professional is not the person's GP, inform the person's GP about the referral. [2009]

19. If a person answers 'yes' to either of the depression identification questions (see recommendation 22) and the practitioner is competent to perform a mental health assessment, review the person's mental state and associated functional, interpersonal and social difficulties. [2009]

20. Consider using a validated measure (for example, for symptoms, functions and/or disability) when assessing a person with suspected depression to inform and evaluate treatment. [2009]

21. If a person has significant language or communication difficulties, (for example people with sensory or cognitive impairments), consider asking a family member or carer about the person's symptoms to identify possible depression. [2004, amended 2017] (See also NICE's guideline on mental health problems in people with learning disabilities.)

22. Conduct a comprehensive assessment that does not rely simply on a symptom count when assessing a person who may have depression. Take into account both the degree of functional impairment and/or disability associated with the possible depression and the length of the episode. [2009]

23. Think about how the factors below may have affected the development, course and severity of a person's depression in addition to assessing symptoms and associated functional impairment:

- any history of depression and coexisting mental health or physical disorders
- any history of mood elevation (to determine if the depression may be part of bipolar disorder)
- any past experience of, and response to, previous treatments
- the quality of interpersonal relationships
- living conditions, employment situation and social isolation. [2009, amended 2017]

Acquired cognitive impairments

24. When assessing a person with suspected depression:

- be aware of any acquired cognitive impairments
- if needed, consult with a relevant specialist when developing treatment plans and strategies. [2009, amended 2017]

25. When providing interventions for people with an acquired cognitive impairment who have a diagnosis of depression:

- if possible, provide the same interventions as for other people with depression
- if needed, adjust the method of delivery or length of the intervention to take account of the disability or impairment. [2009, amended 2017]

Depression with anxiety

26. When depression is accompanied by symptoms of anxiety, the first priority should usually be to treat the depression. When the person has an anxiety disorder and comorbid depression or depressive symptoms, consult NICE guidance for the relevant anxiety disorder if available and consider treating the anxiety disorder first. [2004]

Risk assessment and monitoring

27. Always ask people with depression directly about suicidal ideation and intent. If there is a risk of self-harm or suicide:

- assess whether the person has adequate social support and is aware of sources of help
- arrange help appropriate to the level of need
- advise the person to seek further help if the situation deteriorates. [2004]

28. If a person with depression presents considerable immediate risk to themselves or others, refer them urgently to specialist mental health services. [2004]

29. Advise people with depression of the potential for increased agitation, anxiety and suicidal ideation in the initial stages of treatment. Check if they have any of these symptoms and:

- ensure that the person knows how to seek help promptly
- review the person's treatment if they develop marked and/or prolonged agitation. [2004]

30. Advise a person with depression and their family or carer to be vigilant for mood changes, negativity and hopelessness, and suicidal ideation, and to contact their practitioner if concerned. This is particularly important during high-risk periods, such as starting or changing treatment and at times of increased personal stress. [2004]

31. If a person with depression is assessed to be at risk of suicide:

- take into account toxicity in overdose if an antidepressant is prescribed or the person is taking other medication; (if necessary, limit the amount of medicine available)
- consider increasing the level of support, such as more frequent direct or telephone contacts
- consider referral to specialist mental health services. [2004]

Active monitoring

32. For people who do not want an intervention with less severe depression, in particular those whose depressive symptoms are improving, or people with subthreshold depressive symptoms:

- discuss the presenting problem(s) and any concerns that the person may have
- provide information about the nature and course of depression
- arrange a further assessment, normally within 2 weeks

- make contact if the person does not attend follow-up appointments. [2004]

7 Treatment of new depressive episodes

- Treatment of new depressive episodes: What are the relative benefits and harms of psychological, pharmacological and physical interventions alone or in combination?

7.1 Introduction: Interventions to treat depressive episodes (all severity)

When first choosing an intervention to manage a depressive episode, the clinician and person with depression are faced with a range of treatments. The available range of drug treatments has extended significantly since the introduction of monoamine oxidase inhibitors and tricyclic antidepressants in the 1950s. From the 1980s, selective serotonin reuptake inhibitors were introduced followed by so-called third generation antidepressants such as serotonin and noradrenaline reuptake inhibitors and mirtazapine. Psychological therapies emerged early in the twentieth century with psychoanalytic treatment followed by behavioural, cognitive and interpersonal therapies in the 1950s and 1960s. Recent years have brought incremental developments in psychological interventions and diversification of therapy modalities to include individual, group, long-term, and short-term interventions. Since the early 1990s, there has been an increasing emphasis on improving precision to specifically treat depression (Castonguay and Beutler 2006) and technological advances in recent years have also enabled the development of digital and app-based interventions. Various permutations of combined pharmacological and psychological treatments are possible, extending further the array of interventions for depression. To inform the choice of intervention, knowledge of the relative benefits, harms and costs is essential. It is particularly important to know if combinations of treatments offer any advantages as they likely to be more resource-intensive and more onerous to patients.

This chapter reviews evidence from studies of treatments that are suitable as initial interventions for depression, and evidence is reviewed across a range of pharmacological, psychological and physical interventions in both less and more severe depression. A problem commonly encountered in trying to weigh up a number of interventions is that comparisons between specific interventions that would be informative to patients and clinicians are lacking, particularly between psychological therapies where there is a paucity of head-to-head studies (Farah et al. 2016). Therefore, a network meta-analysis has been conducted as this allows for estimation of comparative effects that have not been investigated head-to-head in randomised clinical trials and ranking of treatment options from best to worst (Caldwell et al. 2005). Network meta-analysis also helps to visualise and interpret the wider picture of the evidence and to understand the relative merits of these multiple interventions to help inform the development of decision aids for patients and clinicians (Mills 2013).

For the purposes of the network meta-analysis, pharmacological treatments have been allocated to three groups: tricyclic antidepressants, selective serotonin reuptake inhibitors and other antidepressants. Psychological therapies are grouped according to common theoretical structure and methodological approach. Older treatments that would no longer be considered clinically suitable (such as the more toxic tricyclic antidepressants) are included in the meta-analysis along with control interventions that would not themselves be of clinical interest, as this maximises the range of comparisons and increases the precision of treatment effect estimates (Caldwell et al. 2005). In depression treatment studies, control interventions are diverse and include pill placebo, attention placebo, and waiting list control. It is known that choice of control condition can influence the apparent effect size of the intervention under investigation with waiting list control generating the largest effect size (Furukawa et al. 2014).

Update 2017

7.1.1 Pharmacological interventions

7.1.1.1 Antidepressants

Selective serotonin re-uptake inhibitors (SSRIs) are by far the most widely prescribed antidepressants and are currently recommended as first-line treatment for moderate to severe depression by most, if not all, authorities (Anderson et al. 2008, NICE 2009, APA 2010). SSRIs are usually well tolerated although nausea, insomnia and agitation can be troublesome at the start of treatment. In the longer term, sexual dysfunction (lowered libido, erectile dysfunction, and delayed orgasm) is fairly common (Fava and Rankin 2002) and hyponatraemia can occur in older people (De Picker et al. 2014). More recently, the effect of SSRIs on platelet aggregation has become better recognised and quantified – risk of bleeding is increased (Jiang et al. 2014), especially when used alongside NSAIDs (Anglin et al. 2014, Oka et al. 2014) aspirin or anticoagulants (Quinn et al. 2014).

SSRIs are fairly safe in overdose (Buckley and McManus 2002) and show little direct cardiac toxicity (Beach et al. 2014), and have minimal effect on cardiac conduction. Two exceptions here are citalopram and escitalopram which prolong QT interval even at clinical doses and show somewhat greater toxicity in overdose (MHRA 2011). However, little evidence has emerged of a substantially increased risk of cardiotoxic events in normal clinical use (Zivin et al. 2013, Qirjazi et al. 2016).

SSRIs have fairly flat dose-response curves in depression and higher doses have not been shown to have greater effect than the minimum effective dose, with the possible exception of sertraline for which doses above 50mg may be more effective (MHRA 2005). Individual SSRIs also differ in their interaction potential, being highest with fluvoxamine, fluoxetine and paroxetine and lowest with citalopram and escitalopram (Hemeryck and Belpaire 2002).

The main alternative to SSRIs is mirtazapine. This is a sedative antidepressant that rarely causes sexual dysfunction or bleeding abnormalities but is associated with weight gain in some people (Watanabe et al. 2011). Its long half-life and strong sedative properties may be problematic at the start of treatment when significant ‘hangover’ is quite common. Trazodone (Brogden et al. 1981) is a broadly similar drug with comparable properties except that weight gain is less likely. Trazodone, although once very widely used, is infrequently prescribed in the UK for depression, although it is a popular sedative in older people.

Venlafaxine, a serotonin and noradrenaline reuptake inhibitor shares many properties with SSRIs (Ellingrod and Perry 1994). It may be slightly more effective but is probably less well tolerated (Smith et al. 2002). It is more toxic in overdose (Buckley and McManus 2002) because of the potential for seizures. Duloxetine is similar to venlafaxine but is probably less toxic in overdose.

Tricyclic antidepressants (TCAs) are still prescribed although they are now not often initiated for depression, at least in primary care. Amitriptyline remains very widely prescribed but much of this prescribing is for pain syndromes and migraine prophylaxis. Nortriptyline is still used in older patients where it is seen as a useful therapeutic agent. Dosulepin (dothiepin) prescribing has fallen dramatically over the past 20 years because of its toxicity in overdose. All TCAs show high overdose toxicity (Cassidy and Henry 1987, Henry et al. 1995) with the exception of lofepramine (which is still used to some extent [Buckley and McManus 1998]) and nortriptyline (Buckley and McManus 2002, Morgan et al. 2004), although some data suggest otherwise in the latter case (Henry et al. 1995).

Since the last guideline, two new antidepressants have come into UK clinical practice. Agomelatine is as effective as other antidepressants and has placebo-level tolerability (Taylor et al. 2014). However, it is a branded drug, unlike all of the antidepressants mentioned so far, and so its purchase cost is relatively high. Concerns over hepatic toxicity have led to the introduction of a monitoring schedule which further limits the drugs utility. Vortioxetine is a multimodal antidepressant as it inhibits the serotonin (also known as 5-

Update 2017

hydroxytryptamine [5-HT]) transporter and modulates 5-HT receptor activity. It is recommended by NICE as a third-line agent for treating major depressive episodes in adults (NICE 2015).

Discontinuation reactions occur with all antidepressants (Taylor et al. 2006) but are particularly marked and frequent with paroxetine and venlafaxine (Schatzberg et al. 2006). Symptoms include insomnia, electric shock sensations, dizziness, mood changes and anxiety. Treatment should always be withdrawn slowly unless a serious adverse event has occurred. A general rule is that the withdrawal should take a few days if the drug has been taken for weeks, a few weeks if taken for months, and a few months if the drug has been taken for years.

Update 2017

The technique of network meta-analysis (NMA) has been used in the literature to assess the comparative efficacy and acceptability of antidepressants. An NMA of modern antidepressants (Cipriani et al. 2009) suggested that sertraline and escitalopram had the best combination of efficacy and tolerability. Mirtazapine and venlafaxine were highly ranked for efficacy only. Reboxetine was ranked last for efficacy and acceptability. A second NMA (Khoo et al. 2015) included fluvoxamine, agomelatine, trazodone and duloxetine which were not examined in the first NMA. Mirtazapine and duloxetine were found to be most efficacious but duloxetine was the least well tolerated. Using numerous outcome measures, agomelatine, mirtazapine and escitalopram showed the best balance of efficacy and acceptability.

7.1.1.21 St John's wort

St John's wort, an extract of the plant Hypericum perforatum, has been used for centuries for medicinal purposes including the treatment of depression. It is not licensed as a medicine in the UK but can be bought 'over the counter' from health food shops, herbalists and community pharmacies. Many different branded preparations are available. St John's wort is licensed in Germany for the treatment of depression.

St John's wort is known to contain at least ten constituents or groups of compo- nents that may contribute to its pharmacological effects (Linde & Mulrow 2004), but its exact mode of action is unknown. These include naphthodianthrons, flavonoids, xanthons and biflavonoids (Wagner and Bladt 1994). In common with all herbal prepa- rations, the quantity and proportions of each constituent varies among batches (Wang et al. 2004). Most commercial products are standardised with respect to hypericin content, but it is not known if this is the only active component. Individual brands or batches of the same brand may, therefore, not be therapeutically equivalent. Many clinically important drug interactions have been reported (Committee on Safety of Medicines 2000). St John's wort may also cause photosensitivity.

7.1.2 Psychological interventions

7.1.2.1 Self-help (without support or with minimal support)

Update 2017

Self-help (without support or with minimal support, also called unguided self-help) 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 can include the provision of information in the form of books or other written materials or audio-recordings that include psychoeducation about the problem and describe techniques to overcome it (for instance, cognitive bibliotherapy and self-examination therapy). Computerised self-administered versions of psychological therapies have also been developed (including computerised-CBT [cCBT] and online positive psychological intervention). A taxonomy has been identified that distinguishes between self-administered work, in which an individual uses the self-help materials exclusively on his or her own (self-help without support), versus minimal contact in which the individual works through the self-help materials with irregular, often non face-to-face contact with a

practitioner whose role is to check on progress and motivate the user (self-help with minimal support), versus self-help with support, see below, in which the individual receives regular and scheduled meetings with a practitioner whose role is to support and guide him or her in using the self-help materials (Glasgow and Rosen 1978).

7.1.2.25 Self-help (with support)

Self-help with support (also called guided self-help) is generally accepted as being more than simply giving people literature to read. Intervention content may overlap with those used in self-help (without or with minimal) support, for instance, cognitive bibliotherapy and computerised psychological therapies (including computerised-CBT [cCBT], computerised psychodynamic therapy, computerised-problem solving therapy and cognitive bias modification), the difference being the regular scheduled support of a healthcare practitioner (for example, PWP) for the purposes of supporting and/or facilitating the individual to complete work with the self-administered materials by introducing, monitoring, and reviewing the outcome of such treatment.

7.1.2.35 Psychoeducational interventions

Psychoeducation is a structured educational treatment (often offered in groups) that provides patients with information about depression, often through a didactic format. 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.

7.1.2.41 Behavioural therapies

Update 2017

Operant or instrumental learning posits that depressive behaviours are learned through the contingencies around those behaviours. In behavioural therapies, depression is seen as the result of a low rate of positive reinforcement and is maintained through negative reinforcement (Ferster 1973). Most commonly, patients use avoidance to minimise negative emotions and situations they worry will be unpleasant in the short-term, which may produce difficulties in the long-term. Behavioural therapies focus on behavioural activation aimed at encouraging the patient to develop more rewarding and task-focused behaviours as well as stepping out of patterns of negative reinforcement. The approach was developed by Lewinsohn (1976) and there are still a group of therapies based on this traditional approach (referred to as behavioural therapy [Lewinsohn 1976] in this guideline). However, more recently there has also been a renewed interest in behavioural activation (for example, Jacobson et al. 2001, Hopko et al. 2003, Dimidjian et al. 2008, Watkins et al. 2011), and it is now known, as a therapy in its own right. There are effectively two strands of behavioural activation. One strand focuses more on increasing positive activities through regular activity scheduling (Hopko et al. 2003). The other strand focuses more on reducing avoidance and understanding a patient's behaviour within his or her particular environment and context. The main approach of the functional-contextual variant of BA is functional analysis, which is the analysis of antecedents, consequences, and variability in behaviour in order to plan effective behavioural change (Jacobson et al. 2001).

Another example of a specific intervention in this category that is linked by a common underlying philosophy is the Coping with Depression (CWD) course most frequently delivered in group format (but also tested in individual format). The CWD course has similarities with psychoeducational group programmes but it was originally developed by Lewinsohn and colleagues (Lewinsohn et al. 1984) and has its roots in social learning theory, according to which depression is associated with a decrease in pleasant and an increase in unpleasant person-environment interactions. Another example of a specific intervention that we have included in this class is social rhythm therapy (SRT) which is based on the theory that disordered circadian biology contributes to the development and maintenance of

depression and that helping patients to develop more regular routines and social patterns will facilitate stabilisation of underlying circadian abnormalities and reduce symptoms.

7.1.2.53 Cognitive and cognitive behavioural therapies

Cognitive behavioural therapy (CBT) for depression was developed by Aaron T. Beck during the 1950s and was formalised into a treatment in the late 1970s (Beck et al. 1979). Its original focus was on the styles of conscious thinking and reasoning of depressed people, which Beck posited was the result of the operation of underlying cognitive schemas or beliefs. The cognitive model describes how, when depressed, people focus on negative views of themselves, the world, and the future. The therapy takes an educative approach where, through collaboration, the person with depression learns to recognise his or her negative thinking patterns and to re-evaluate his or her thinking. This approach also requires people to practise re-evaluating their thoughts and new behaviours (called homework). The approach does not focus on unconscious conflicts, transference, or offer interpretation as in psychodynamic psychotherapy. There is also an important emphasis on increasing activity and engaging in rewarding behaviours, as per behavioural activation, as well as the use of behavioural experiments to test underlying beliefs. As with any psychological treatment, cognitive behavioural therapy is not static and has been evolving, and in addition to the continued individual-format high-intensity CBT, CBT has also been delivered in a group format and in a low-intensity format. This guideline used the cut-off of 15 sessions to distinguish between a longer course of CBT (over 15 sessions) and briefer courses of CBT (under 15 sessions).

The principles of CBT also form the basis of a number of other stand-alone interventions that are grouped under this class, including, problem solving. Problem solving interventions are based on the theory that depression is associated with social problem-solving difficulties (Nezu 1987) which may relate to the effects of the depressed state, lack of knowledge, and/or rumination (Watkins 2008) and aims to help patients solve problems and develop problem-solving skills (Nezu et al. 1989) in order to improve depression symptoms. Also drawing on common cognitive and cognitive behavioural principles although with a different emphasis and with some different techniques are a newer wave or so-called third wave of cognitive therapies including acceptance and commitment therapy (ACT) and mindfulness-based cognitive therapy (MBCT). These therapies encourage mindfulness of internal experiences and emphasize acceptance instead of change of negative internal sensations and thoughts (Herbert et al. 2009). Another, albeit older, variant of the traditional Beckian cognitive behavioural approach is rational emotive behaviour therapy (REBT) which was developed by Ellis in the 1950s (Ellis 1955), and which proponents believe may promote a deeper change through advocating unconditional self-acceptance, focusing explicitly on reducing secondary problems such as depression about depression (meta-emotions) and explicitly targeting demandingness (imperative or absolutistic demands on self, others, and life), the latter of which is considered the crucial component of depression.

Update 2017

7.1.2.5.4 Mindfulness-based cognitive therapy

Mindfulness-based cognitive therapy (MBCT) was developed with a specific focus on preventing relapse/recurrence of depression (Segal et al. 2002, Kuyken et al. 2008, Kuyken et al. 2015) which is covered in Chapter 11. It is an 8-week manualised group-based skills training programme with each session lasting 2 hours, and four follow-up sessions in the year after the end of therapy. It integrates the use of mindfulness mediation as derived from mindfulness-based stress reduction (Kabat-Zinn 1990), with psychoeducation and principles from CBT for acute depression (Beck et al. 1979). 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). Through guided meditative practice, 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 'decentre' from negative thoughts and feelings, accept difficulties using a stance of self-compassion and use bodily awareness to ground and transform experience. Patients develop an 'action plan' that sets out strategies for responding when they become aware of early warning signs of relapse/recurrence.

7.1.2.5.25 Rumination-focused cognitive behavioural therapy

Rumination-focused cognitive behavioural therapy (RFCBT) was developed to specifically target rumination, (repetitive negative thinking about the causes, meanings, and implications of symptoms, problems and upsetting events), which has been robustly identified as an important contributory factor to the onset and maintenance of depression and other disorders (Nolen-Hoeksema et al. 2008). Rumination is a common residual symptom of depression and associated with poor recovery. RFCBT was therefore designed and evaluated for severe, chronic and residual depression (Watkins et al. 2011, Hvennegard et al. 2015, Teismann et al. 2014). It is a manualised treatment deliverable in individual, group and internet formats (Watkins 2016). Based on evidence that rumination is a mental habit (Watkins and Nolen-Hoeksema 2014), patients learn to notice warning signs for rumination, and establish alternative adaptive coping behaviours, through functional analysis and repeated practice. Based on theory and evidence that thinking style determines whether repetitive thinking has helpful versus unhelpful consequences (Watkins, 2008), these strategies focus on shifting thinking style including exercises to increase concrete and specific thinking, absorption in positive activities, and self-compassion, rather than directly challenging negative thoughts.

7.1.2.5.32 Cognitive Behavioural Analysis System of Psychotherapy

Cognitive Behavioural Analysis System of Psychotherapy (CBASP) is a variant of CBT designed solely and specifically to treat chronic depression (McCullough 2003) which is covered in Chapter 9. CBASP is based on the theoretical view that patients with chronic depression have become disconnected from their environment and thus are not able to change their behaviour or learn in response to environmental feedback, which has negative consequences especially for interpersonal relationships. It differs from standard CBT by an increased emphasis on directing the patient's attention to the effect of his or her actions on others, including the therapist, through a technique called Situational Analysis that explores in detail sequences of events, actions, and consequences. In addition, patients are encouraged to increase empathic behaviour to others, and the therapist uses his or her own responses to reduce unhelpful in-session behaviours from the patient. CBASP has predominantly been examined in the context of chronic depression (lasting more than 2 years), and combined with antidepressant medication (Keller et al. 2000, Klein et al. 2004, Schramm et al. 2011, Wiersma et al. 2014).

Update 2017

7.1.2.67 Counselling

Counselling was developed by Carl Rogers (1957) who believed that people had the means for self-healing, problem resolution and growth if the right conditions could be created. These conditions include the provision of positive regard, genuineness and empathy. Rogers's original model was developed into structured counselling approaches by Truax and Carkhuff (1967) and, independently, by Egan (1990) who developed the three stage model: exploration, personalizing, and action. Voluntary sector counselling training (for example, Relate) tends to draw on these models. However, although many other therapies now use the basic ingredients of client-centred counselling (Roth and Fonagy 2005), there are differences in how they are used, for instance, emotion-focused therapy (EFT) and relational client-centered therapy. A more directive form of counselling has also developed, that incorporates elements of supportive listening and history taking in common with non-directive counselling but also includes more directive techniques of problem clarification, goal formation and problem solving. Counselling has become a generic term used to describe a broad range of interventions delivered by counsellors usually working in primary care. The

content of these various approaches may include psychodynamic, systemic or cognitive behavioural elements (Bower et al. 2003).

7.1.2.7 Interpersonal psychotherapy

Interpersonal therapy (IPT) was developed by Klerman and Weissman (Klerman et al. 1984) initially for depression although it has now been extended to other disorders (Weissman et al. 2000). IPT focuses on current relationships, not past ones, and on interpersonal processes rather than intra-psychic ones (such as negative core beliefs or automatic thoughts as in CBT, or unconscious conflicts as in psychodynamic psychotherapy). It is time limited and focused on difficulties arising in the daily experience of maintaining relationships and resolving difficulties during an episode of major depression. Early in the treatment, patient and therapist agree to work on a particular focal area that would include: interpersonal role transitions, interpersonal roles/conflicts, grief and/or interpersonal deficits. IPT is appropriate when a person has a key area of difficulty that is specified by the treatment (for example, grief or interpersonal conflicts). It can be delivered as an individually focused therapy but has also been developed as a group therapy (Wilfley et al. 2000). The character of the therapy sessions is, largely, facilitating understanding of recent events in interpersonal terms and exploring alternative ways of handling interpersonal situations. Although there is not an explicit emphasis on 'homework', there is an emphasis on effecting changes in interpersonal relationships and tasks towards this end may be undertaken between sessions.

7.1.2.8 Short-term psychodynamic psychotherapies

Short-term psychodynamic psychotherapies are based on psychoanalytic techniques but may often be considerably briefer than psychoanalysis proper. Short-term psychodynamic psychotherapy considers the symptoms of depression as the result of core relationship conflicts predominately based on early experience and aims to help the person become aware of the link between conflicts and symptoms using the therapeutic relationship as a central vehicle for insight and change. As with other schools of psychological therapy, there are a number of variations on the original model of psychodynamic psychotherapy. Some approaches focus on the dynamic of drives (for example, aggression) while others focus on relationships (Greenberg and Mitchell 1983). Other forms of this therapy have been influenced by attachment theory (Holmes 2001). Clinical trials of psychodynamic psychotherapy have traditionally focused on short-term psychological therapy (typically 10 to 30 weeks) usually in comparison with antidepressants or CBT.

7.1.2.9 Long-term psychodynamic psychotherapies

A number of recent trials have examined a longer-term version of psychodynamic psychotherapy with treatment durations of up to three years. Long-term psychodynamic psychotherapy is an intensive, transference-based therapeutic approach and acts in a supportive-interpretive continuum (depending on the therapeutic needs of the patient) in order to explore and work through a broad range of intrapsychic and interpersonal conflicts (Gabbard 2004).

7.1.2.10 Behavioural couples therapy

Therapists have noted that a partner's critical behaviour may trigger an episode of depression, and/or maintain or exacerbate relapse in the long term (for example, Hooley and Teasdale 1989), although other researchers have questioned this (for example, Hayhurst et al. 1997). There has also been some research looking at differences in the vulnerabilities between men and women within an intimate relationship, with physical aggression by a partner predicting depression in women. Difficulties in developing intimacy, and coping with conflict, also predict depression in both men and women (Christian et al. 1994). Couples therapy has evolved in recent years. Systemic couples therapy aims to give the couple new

Update 2017

perspectives on the presenting problem (for example, depressogenic behaviours), and explore new ways of relating (Jones and Asen 1999). Other developments such as those by Jacobson and colleagues (1993) took a more behavioural approach. In the analysis of behavioural couples therapy in this guideline, the focus of the search was not on a specific approach but on couples therapy more generally.

7.1.3 Psychosocial interventions

Psychosocial interventions are non-pharmacological and address psychological aspects in a broader societal or familial perspective. An example of a group of psychosocial interventions for depression include peer-mediated support. 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 have a history of depression themselves are recruited and trained to deliver interventions. These interventions can include befriending and mentoring. Befriending can also include volunteers without a history of depression. 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. However, the primary goal of these interventions is to enable mutual support by bringing people with depression into contact with other people who are having similar experiences and providing opportunities for sharing problems and solutions.

7.1.4 Physical interventions

7.1.4.1 Electroconvulsive therapy (ECT)

Update 2017

Electroconvulsive therapy (ECT) has been used as a treatment for depression since the 1930s. In its modern form ECT is perceived by many healthcare professionals to be a safe and effective treatment for severe depression that has not responded to other standard treatments (Geddes et al., 2003b). But many others, including some patient groups, consider it to be an outdated and potentially damaging treatment (Rose et al., 2003). During ECT, an electric current is passed briefly through the brain, via electrodes applied to the scalp, to induce generalised seizure activity. The therapeutic effects of seizure induction may arise from changes in cerebral blood flow and metabolism or subsequent effects on nerve growth, neurotransmitter pathways, and neuroendocrine systems (Anderson and Fergusson, 2013).

The person receiving treatment is placed under general anaesthetic and muscle relaxants are given to prevent body spasms. The ECT electrodes can be placed on both sides of the head (bilateral placement) or on one side of the head (unilateral placement). Unilateral placement is usually to the non-dominant side of the brain, with the aim of reducing cognitive side effects. The standard bilateral placement is bitemporal/temporofrontal but some studies have used bifrontal placement in the hope of reducing cognitive side effects associated with the standard placement. Electro-encephalogram (EEG) monitoring of ECT treatment and the use of shorter electrical pulse appear to limit cognitive side-effects and there is now interest in the use of even shorter (ultra-brief) pulses (Tor et al. 2015). The number of sessions undertaken during a course of ECT usually ranges from six to twelve, although a substantial minority of patients respond to fewer than six sessions. ECT is usually given twice a week in the UK; less commonly it is given once a fortnight or once a month as continuation or maintenance therapy to prevent the relapse of symptoms. It can be given on either an inpatient or day patient basis.

ECT causes short-term disorientation immediately after treatment and may cause short- or long-term memory impairment for past events (retrograde amnesia) and current events (anterograde amnesia). These effects appear to be dose related and depend on electrode placement, possibly the type of electrical stimulus and patient characteristics (Ingram et al.

2008). However the persistence, severity and precise characterisation of such impairments are still a subject of debate. There is some evidence that prolonged short-term disorientation immediately after treatment predicts retrograde amnesia after the end of a course of treatment (Sobin et al. 1995) but not two months after the course. Cognitive impairments have been highlighted as a particular concern by many patients, especially retrograde amnesia for autobiographical events (Rose et al., 2003). There is no simple relationship between subjective cognitive impairment and cognitive test measures, which has contributed to the polarisation of views about the relative risks and benefits of ECT. At present there is a lack of consensus as to the best method of assessing cognitive function during a course of ECT. The benefit of using only a global measure such as the mini-mental state examination in its original or modified form (3MSE) is uncertain given the inconsistent effects of ECT on these measures in trials. And given the evidence that the ability to learn new material (anterograde memory) recovers after the end of ECT treatment, a main concern is in the early detection and minimisation of persistent retrograde memory loss, particularly for important autobiographical memories. Detecting cognitive impairments only at the end of treatment does not give the practitioner the opportunity to alter treatment to attempt to minimise this, although it may lead the practitioner to consider cognitive remediation; there is no evidence, however, to show that this is effective. A battery consisting of a formal mood rating scale (MADRS), the 3MSE, an autobiographical memory task, a word learning task, and tests of digit span forward and backward has been suggested (Porter et al., 2008), but it takes an hour to administer.

In line with NICE policy regarding the relationship of technology appraisals to clinical practice guidelines, this guideline updates the NICE technology appraisal guidance on the use of electroconvulsive therapy (TA59) only for depression in adults (the TA covers the use of ECT in the treatment of mania and schizophrenia as well as depression in children and adolescents; NICE 2003).

Update 2017

Key points to emerge from the reviews underpinning the NICE TA on ECT (NICE 2003), which concluded that ECT is an effective treatment, include:

- real ECT had greater short-term benefit than sham ECT
- ECT had greater benefit than the use of certain antidepressants
- bilateral ECT was reported to be more effective than unilateral ECT
- the combination of ECT with pharmacotherapy was not shown to have greater short-term benefit than ECT alone
- cognitive impairment does occur but may only be short term
- compared with placebo, continuation pharmacotherapy with tricyclic antidepressants and/or lithium reduced the rate of relapses in people who had responded to ECT
- preliminary studies indicate that ECT is more effective than repetitive transcranial magnetic stimulation.

In the 2009 update of this Guideline, it was observed that maintenance ECT is used on a small scale in the United Kingdom for people with recurrent depression that is not responsive to other treatments but with considerable uncertainty about its long-term efficacy, acceptability, and possible side-effects (including cognitive impairment), The Guideline concluded, therefore, that further studies were required of the effectiveness of maintenance ECT for relapse prevention in people with severe and recurring depression that does not respond to pharmacotherapy or psychological treatment.

7.1.4.26 Exercise

The effect of physical activity on mental health has been the subject of research for several decades. There is a growing body of literature examining the effects of physical activity in the treatment of depression. The aerobic forms of physical activity, especially jogging or running, have been most frequently investigated. In recent years ‘exercise on prescription’ schemes

have become popular in primary care in the UK (Biddle et al.1994), many of which include depression as a referral criterion.

Guidelines for physical activity referral schemes have been laid down by the Department of Health (2001, Mead et al. 2008). Several plausible mechanisms for how physical activity affects depression have been proposed. In the developed world, regular physical activity is seen as a virtue; the depressed patient who takes regular physical activity may, as a result, get positive feedback from other people and an increased sense of self-worth. Physical activity may act as a diversion from negative thoughts and the mastery of a new skill may be important (Lepore 1997; Mynors-Wallis et al. 2000). Social contact may be an important benefit, and physical activity may have physiological effects such as changes in endorphin and monoamine concentrations (Thoren et al.1990; Leith1994).

Update 2017

For the purposes of the guideline, physical activity is defined as a structured physical activity with a recommended frequency, intensity and duration when used as a treatment for depression. It can be undertaken individually or in a group. Physical activity may be divided into aerobic forms (training of cardio-respiratory capacity) and anaerobic forms (training of muscular strength/endurance and flexibility/co-ordination/relaxation) (American College of Sports Medicine, 1980). In addition to the type of physical activity, the frequency, duration and intensity should be described. Within the network meta-analysis, interventions based on structured physical activity have been grouped with yoga-based interventions.

Yoga is a method based on traditional Indian philosophical and spiritual practices with modern yoga forms used in the western world being mostly associated with physical postures, breathing techniques, and meditation. Yoga is advocated for people living with chronic pain or physical illness; a recent systematic review reported a small number of inclusive studies of yoga in the treatment of depression (Cramer 2017).

7.1.4.3 Light therapy

Depression with a seasonal pattern as a separate diagnosis has been less accepted in Europe than North America, and an alternative view is that major depression with a seasonal pattern is an extreme form of a dimensional 'seasonality trait' rather than a specific diagnosis with so-called 'subsyndromal major depression with a seasonal pattern' appearing to be common. Nevertheless there are some patients with recurrent major depression who experience a seasonal pattern to their illness, at least for a time. There also appear to be people who experience seasonal fluctuations in mood that do not reach criteria for major depression.

The hypothesis that light therapy (that is, increasing the amount or duration of light exposure) might be an effective treatment is based on the presumption that depression with a seasonal pattern is caused by a lack of light in the winter months; its benefit may be due to its effects on built-in circadian rhythms (Lewy et al. 1987). In light therapy, a box of fluorescent tubes is used to provide light of specific intensity and duration.

The 2009 guideline concluded that, due to the small number of inconclusive trials, further trials of adequate size were necessary to evaluate the efficacy of light therapy compared with antidepressant medication for mild to moderate depression with a seasonal pattern.

7.1.4.4 Acupuncture

The medical use of acupuncture combines theoretical principles of traditional Chinese medicine, such as re-balancing bodily energy, with knowledge of physiology and anatomy to determine the appropriate site of application. There are several styles of treatment including classical, auricular, trigger point and single point acupuncture. Variations on the traditional insertion of needles include electro-acupuncture and laser acupuncture (Smith CA et al. 2010). It has been suggested that the therapeutic effects of acupuncture may be mediated by its action on limbic brain structures, including the cingulate cortex (Napadow et al. 2005).

Update 2017

Acupuncture may be used as a stand-alone intervention or in combination with antidepressant treatment (Chan et al. 2015). Minor side-effects include bleeding and pain at the needling site. The risk of serious adverse effects is reported to be low; they include nerve trauma, pneumothorax, infection at the puncture site and transmission of hepatitis B (White et al. 2004).

7.1.56 Combined interventions

Evidence indicates that only one in three people reach remission using first-line antidepressant monotherapy and in these cases only after a typical delay of 6 weeks or more (Trivedi et al., 2006). Partly in response, clinical trials have investigated whether the co-initiation of two or more treatments might produce a greater or more accelerated treatment effect. Biological co-initiation trials have investigated pharmaceutical-pharmaceutical or pharmaceutical-nutraceutical (pharmaceutical grade, standardised nutrient) combinations, aiming at rational strategies with complimentary modes of central nervous system activity and low risk of interaction. These trials have included the co-initiation of mirtazapine with each of fluoxetine, venlafaxine or bupropion (against fluoxetine monotherapy, Blier et al. 2009); sertraline co-initiated with triiodothyronine (T3) (Cooper-Kazaz et al. 2007); SSRIs with pindolol (Ballesteros and Callado 2004) or with omega-3 fatty acids (for example Gertsik et al. 2012). However, the treatment duration of these trials is limited (typically 4 – 8 weeks) making it difficult to fully assess effects, including harmful effects, and since remission can be achieved with single agent antidepressants, these immediate combination strategies risk exposing patients to unnecessary additional side effects, expense or physical monitoring, and taking medicines that are not licensed for use in depression. For these reasons immediate biological combinations are currently difficult to justify as first-line treatment for depression (Rush 2010).

A broader alternative comes through the combination of different treatment modalities, for example through biological-psychological or biological-exercise strategies. Contemporary neuroscience provides an understanding of how conscious psychological work (processed proximally by evolved prefrontal areas of the brain) may naturally integrate with pre-conscious antidepressant effects (working proximally at the limbic level, for example Fu et al. 2004, Norbury et al. 2009). Evidence-based theories of antidepressant action also now highlight the importance of social and physical activity in mediating the initial neuropsychological effects of antidepressants (Pringle and Harmer 2015). Considered this way, antidepressants, exercise and psychological interventions offer potentially complimentary ways to treat depression. Where the formulation includes both biological vulnerability to depression and psychological maintaining factors then an initial combined approach (through antidepressants and psychological interventions) may simply offer the most powerful intervention, where this is acceptable and available. Alternatively, medication that restores sleep, motivation or cognitive ability may enable fuller, more effective use of psychological or exercise interventions.

Alongside the potential treatment benefits of immediate cross-modality combinations there should be some consideration of potential harms. For example, where medication is initiated alongside exercise programmes, the acute pharmacological effects of antidepressants (including possible postural hypotension) and loss of muscular conditioning after periods of inactivity should be considered. Given the availability of monotherapy, the potential harms and limits of our understanding in this area should be discussed with the patient prior to immediate co-initiation strategies.

7.27 Categorisation of the study population according to the symptom severity of the new depressive episode

According to their baseline level of depressive symptom severity, two study populations were identified: people with a new episode of less severe depression and people with a new

Update 2017

episode of more severe depression. These two populations were considered separately, in 2 distinct review questions.

The GC were aware that in order to undertake an NMA, the population included in the analysis should be relatively homogenous; significant differences in the nature or severity of the depressive disorders in the trial populations and their impact as moderators of treatment effect could invalidate the analysis, The GC considered a number of factors which might impact on treatment outcomes such as chronicity or treatment resistance but these were already addressed under separate review questions. The GC also considered whether different types of depression such as melancholia or atypical depression might also respond differently to treatment but work on previous NICE guidelines (for example NICE 2009) and more recent analyses did not support such an approach (for example Cuijpers et al. 2017). The GC considered that treatment severity was a factor which could moderate treatment effects. Symptom severity has long been considered a potential mediator of treatment effect (Sotsky et al. 1991) both within treatments, (Fournier et al. [2010] showed that antidepressant response in relation to placebo varied in clinical importance with severity) and between treatments (for example between CBT and antidepressants DeRubeis et al. [2014]). More recent studies have suggested that difference between treatments may not be so marked, for example Weitz et al. (2017) suggested no difference in response by severity for either CBT or antidepressants, but it should be noted that in the population severity rating on the HRSD 17% would be rated as severe by the criteria adopted by this guideline, in contrast the baseline rating from the BDI indicated that almost 50% would be in the severe range. The GC were also concerned that certain interventions, for example self-help with support were typically only provided to participants with less severe depression and here there was evidence of an impact of severity on outcomes (Button et al. 2013). Having taken these factors into consideration the GC therefore decided that having 2 separate networks for more and less severe depression was the right approach to take.

For a number of interventions specifically behavioural couples therapy, nortriptyline in older people, acupuncture, omega fatty acids and peer support the GC were concerned that the populations in these interventions may differ from the general population in both networks and so separate pairwise comparisons were undertaken for those groups. In order to explore general outcomes of older people and whether there were differences in outcomes for inpatients and community populations, sub-group analyses of the NMA data were undertaken.

The level of severity of the new depressive episode in participants in each RCT was determined by their mean baseline score on one of the depressive symptom scales of those considered in the clinical data analysis. A hierarchy of selected scales was used to prioritise data for extraction; this hierarchy also determined the scale used to estimate the baseline symptom severity of participants in each RCT, if baseline data on more than one depressive symptom scales were reported.

Categorisation of the population in each RCT into one of the two depressive symptom severity levels (that is, less severe and more severe depression) and, consequently, into one of the two review questions was based on an estimated cut-off point on the depressive symptom scale reported in the study. If the mean baseline symptom score of study participants was below the cut-off point, the study was allocated to the review question for people with less severe depression; otherwise, the study was included in the review question for people with more severe depression.

Where information on the baseline mean symptom scale score was not available in a study, studies were categorised according to inclusion criteria, read-outs from figures where these were available, or in rare cases according to the author's description. This option was only used where no other option was available and we were confident that the author's description was likely to be accurate, i.e. where a population was described as mild we were confident that they would be in the less severe category, however greater caution was

Update 2017

exercised in papers describing themselves as moderate or severe due to the location of our cut-off point. When no information was available on the baseline symptom severity of the population included in the RCT, this RCT was excluded from further consideration.

7.2.14 Method for determining cut-off scores for less and more severe depression on each depression scale

In the development of the NMA the GC considered that the severity of depression was a potentially important moderator in determining the outcome of depression treatment. This was based on a number of previous reviews (NICE 2009, Fournier et al. 2010), that suggested that initial severity impacted on recovery and that different treatments might have differential clinical and cost-effectiveness depending on severity (Simon et al. 2006). The commonly used categorisation of depression severity includes persistent sub-threshold symptoms (also known as dysthymia) and mild, moderate and severe depression. The GC considered what would be the most useful division of depression severity on which to base recommendations and decided on a distinction between less severe depression (including subthreshold symptoms) and more severe depression. The GC decided on this distinction because they agreed that it would be most useful in guiding clinical decisions and therefore in the construction of recommendations. The distinction is very similar to that adopted by the NICE Depression guideline (NICE 2009) which primarily used the terms mild to moderate depression and moderate to severe depression when drawing up recommendations.

Having made this decision there was a need to develop a robust and reliable method of classifying studies into these categories. Unfortunately, there is no agreed, commonly used system for classifying depression that is used routinely in clinical trials of depression and which could inform the classification of depression severity. Indeed, a number of studies do not use any such classificatory systems with a diagnosis of depression being the main entry requirement for a trial, others might use terms such as treatment-resistant depression or chronic depression but this does not always relate directly to severity.

The most straightforward way to address this problem is to use the score at entry to a trial of the commonly used standard outcome measures (see below) as an indicator of severity, as these scores are reported in almost all trials:

- MADRS (Montgomery Åsberg Depression Rating Scale)
- HAMD (Hamilton Depression Rating Scale)
- QIDS (Quick Inventory of Depressive Symptomatology)
- PHQ-9 (Patient Health Questionnaire 9 items)
- CES-D (Center for Epidemiologic Studies Depression Scale Revised)
- BDI (Beck Depression Inventory) version I or II.
- HADS-D (Hospital Anxiety and Depression Scale - depression subscale)
- HADS (Hospital Anxiety and Depression Scale - full scale).

However, when this approach was considered, further problems were encountered; first not all commonly used measures report cut-offs for severity (for example the CES-D reports no distinction between mild, moderate or severe); secondly, where they are reported a consistent cut-off is not always used (for example different cut-offs for caseness in the MADRS are reported) and thirdly, the classificatory system was not consistent with the approach adopted by the GC (for example the PHQ-9 which refers to subthreshold symptoms [below caseness] as mild depression). In addition, a review of the relevant literature identified no substantial body of work that allowed for a 'read-across' between scales, although some work has been published on a limited number of scales (for example, Cameron et al. 2008).

Update 2017

In the absence of a substantial literature base to inform the classification of depression, the GC developed a practical approach to determining appropriate cut-offs for more and less severe depression. In doing so the following steps were taken:

- The trials were reviewed and all scales that were used in those trials were identified, relevant papers and manuals which supplied data on caseness thresholds and rating of severity were identified and reviewed.
- The caseness thresholds for all scales were identified as well as the maximum score that was possible to obtain on each scale.
- The content of each scale was then reviewed and an estimation of the degree of 'redundancy' in each scale was made. This was necessary as depression rating scales typically cover a range of different symptom 'clusters' including cognitive, somatic, anxiety and mood, not all of which may be present in an individual with a diagnosis of depression but all of which do need to be present in a rating scale. This results in a necessary 'redundancy' in all depression scales which needs to be taken into account when estimating severity by scores on a scale. This meant that an approach which simply took the value for caseness and the maximum score of the scale could be misleading, depending on the degree of redundancy in a scale. This problem is further complicated by the fact that the commonly used measures vary considerably with a maximum score obtainable from 21 on the HADS to 63 on the BDI-II. To address this problem all scales were carefully reviewed and an estimation of the degree of redundancy (r) was made and checked with the GC. These estimates are listed in Table 42 and were used to determine an 'estimated' cut-off score for severe depression (esd), by applying an estimate of redundancy (r) for each scale to the difference between the maximum score on the scale (m) and the threshold for caseness(c).
- The distinction point (dp) between more and less severe was calculated by dividing the difference between esd and c by 2 and then adding that to c. It is expressed in the equation given below. Where calculations did not result in a whole number, as a general approach numbers were rounded up or down according to standard procedures but some adjustments were made in particular for those scales with a lower total score (that is the HADS, the PHQ-9 and the QIDS-10).

$$dp = \frac{(m-c)(1-r)}{2} + c$$

- The output of this procedure was checked with the GC and also compared with the rating of severity for those scales which had published severity levels. Broadly there was good agreement (a difference of one or two points in most cases) except for the PHQ-9 (see comment above). The cut-offs also had some external validity, for example the cut-off on the HAMD of 24 was very similar to the point at which antidepressant drugs separated from placebo in terms of clinical importance in the meta-analysis by Fournier et al (2010) which is held to be an important distinction between more and less severe depression.

The details of all relevant scales and the agreed distinction point are given in Table 42.

Table 42: Depressive symptom scale characteristics and cut-off points used to determine less severe and more severe depression

	Number of items	Range of scores	Caseness threshold (^{1}c)	Less Severe range	More severe range
MADRS (r= 0.4)	10	0-60	11	11-26	27+
HAMD (17) (r= 0.4)	17	0-60	8	8-23	24+
QIDS-10 (r= 0.2)	10	0-27	6	6-16	17+

Update 2017

	Number of items	Range of scores	Caseness threshold ([1]c)	Less Severe range	More severe range
PHQ-9 (r= 0.2)	9	0-27	10	10-17	18+
CES-D (r= 0.4)	20	0-60	16	16-28	29+
BDI- I ([2]r= 0.5)	21	0-63	12	12-24	25+
BDI- II (r= 0.5)	21	0-63	14	14-26	27+
HADS (r= 0.2)	7	0-21	8	8-15	16+
Notes: [1]c = threshold for caseness [2]r = the redundancy constant					

Although CGI-I (Clinical Global Impressions – Improvement scale) data were considered in relation to the dichotomous outcome of response, as described in section 7.3.4, continuous data based on the CGI-I or the CGI-S (Clinical Global Impression – Severity Scale) were not extracted, and CGI scores were not used to estimate baseline symptom severity, as this was not considered appropriate.

7.36 Methods for clinical evidence synthesis

Update 2017

7.3.17 Network meta-analytic techniques - introduction

Network meta-analytic techniques were employed to synthesise evidence on pharmacological, psychological, combined and physical interventions and estimate the comparative effectiveness between all pairs of interventions considered in each review question covered in this chapter. Network meta-analysis (NMA) takes all trial information into consideration, without ignoring part of the evidence and without introducing bias by breaking the rules of randomisation (for example, by making “naive” addition of data across relevant treatment arms from all RCTs). NMA is a generalization of standard pairwise meta-analysis for A versus B trials, to data structures that include, for example, A versus B, B versus C, and A versus C trials (Dias et al., 2011; Lu & Ades, 2004). A basic assumption of NMA methods is that direct and indirect evidence estimate the same parameter, that is, the relative effect between A and B measured directly from an A versus B trial, is the same with the relative effect between A and B estimated indirectly from A versus C and B versus C trials. NMA techniques strengthen inference concerning the relative effect of two treatments by including both direct and indirect comparisons between treatments, and, at the same time, allow simultaneous inference on all treatments examined in the pair-wise trial comparisons, which is essential for consideration of treatment in economic analysis (Caldwell et al., 2005; Lu & Ades, 2004). Simultaneous inference on the relative effect a number of treatments is possible provided that treatments participate in a single “network of evidence”, that is, every treatment is linked to at least one of the other treatments under assessment through direct or indirect comparisons.

A key assumption when conducting NMA is that the populations included in all RCTs considered in the NMA are similar so that the treatment effects are exchangeable across all populations (Mavridis et al., 2015).

Although the vast majority of RCTs included in the guideline systematic reviews covered in this chapter were considered to have study populations that were similar enough to allow inclusion of RCTs in the NMA, the study populations in a number of RCTs were considered

to differ, and therefore these studies were analysed separately, via pairwise meta-analysis. Details of these studies and the reasons for considering them separately are provided in relevant sections of this chapter.

Full details on the methods used in the NMAs conducted for each review question covered in this chapter are reported in Chapter 17. An overview of included populations, interventions, outcomes and NMA methods is provided in the sections that follow.

7.3.27 Populations, interventions and classes considered in the NMAs

Separate NMAs were conducted for adults with a new episode of less severe depression and adults with a new episode of more severe depression, as defined in Section 7.2.

The following classes and interventions were considered as part of the decision problem, i.e. as interventions considered for recommendation for this review question, in each NMA, according to the availability of respective evidence for each population and on each outcome considered:

Pharmacological interventions

- Class of SSRIs: citalopram, escitalopram, fluoxetine, sertraline
- Class TCAs: amitriptyline, lofepramine
- Mirtazapine (comprising its own class)

Psychological interventions

- Class of self-help (without or with minimal support): cognitive bibliotherapy, computerised-CBT (cCBT), online positive psychological intervention, self-examination therapy
- Class of self-help with support: cognitive bibliotherapy with support, cognitive bias modification with support, computerised psychodynamic therapy with support, computerised-CBT (cCBT) with support, computerised-problem solving therapy with support, tailored computerised-CBT (cCBT) with support
- Class of psychoeducational interventions: psychoeducational group programme
- Class of behavioural therapies: behavioural activation (BA), behavioural therapy (Lewinsohn 1976), coping with depression course (individual), coping with depression course (group), social rhythm therapy (SRT)
- Class of cognitive and cognitive behavioural therapies: CBT individual (under 15 sessions), CBT individual (over 15 sessions), CBT group (under 15 sessions), CBT group (over 15 sessions), problem solving, rational emotive behaviour therapy (REBT), third-wave cognitive therapy individual, third-wave cognitive therapy group
- Class of counselling: directive counselling, emotion-focused therapy (EFT), non-directive counselling, relational client-centered therapy, counselling (any type)
- Class of interpersonal psychotherapy (IPT): IPT
- Class of short-term psychodynamic psychotherapies: psychodynamic counselling, short-term psychodynamic psychotherapy group, short-term psychodynamic psychotherapy individual
- Class of long-term psychodynamic psychotherapies: long-term psychodynamic psychotherapy individual

Physical interventions

- Class of exercise: exercise, yoga

Update 2017

Combined interventions

- Class of combined cognitive and cognitive behavioural therapies with antidepressant: interventions include any intervention belonging to the cognitive and cognitive behavioural therapies class combined with any of the antidepressants considered in the NMAs as part of the decision problem
- Class of combined counselling with antidepressant: interventions include any intervention belonging to the counselling class combined with any of the antidepressants considered in the NMAs as part of the decision problem
- Class of combined IPT with antidepressant: interventions include IPT (i.e. the only intervention within the IPT class) combined with any of the antidepressants considered in the NMAs as part of the decision problem
- Class of combined short-term psychodynamic psychotherapies with antidepressant: interventions include any intervention belonging to the short-term psychodynamic psychotherapies class combined with any of the antidepressants considered in the NMAs as part of the decision problem
- Class of combined long-term psychodynamic psychotherapies class with antidepressant: interventions include any intervention belonging to the long-term psychodynamic psychotherapies class combined with any of the antidepressants considered in the NMAs as part of the decision problem
- Class of combined exercise with antidepressant or CBT: interventions include any intervention belonging to the exercise class with any of the antidepressants considered in the NMAs as part of the decision problem or with any intervention belonging to the cognitive and cognitive behavioural therapies class

The following controls were included in the analysis:

- Pill placebo
- Attention placebo
- Treatment as usual (TAU) class, including TAU and enhanced TAU
- Wait list

In addition to the above interventions and classes, a number of other interventions were included in the NMAs without being part of the decision problem, in order to provide links between interventions of interest and allow indirect comparisons between them:

Imipramine, which belongs to the TCA class, was not part of the decision problem. However, it was included in the clinical analysis because it has been used as a comparator in many drug trials, and therefore comprised a link that allowed indirect comparisons between interventions of interest.

Combined psychological interventions plus pill placebo were retained in the NMA in order to provide links between psychological and/or combined interventions of interest. These interventions were included in a separate class of psychological intervention plus pill placebo.

A number of RCTs that assessed interventions that were not directly part of the decision problem were included in the NMAs. This inclusion was necessary in order to:

- Connect otherwise unconnected networks, so that the relative outcomes between all pairs of interventions considered in each NMA were possible to estimate
- Increase the available evidence on combined interventions and classes, as there was very limited evidence on combination therapies that formed part of the decision problem.

The following studies were included in the appropriate network (for less severe and more severe depression):

Update 2017

1. Studies that included arms of a 'TCA' (comprising a mixture of more than one TCAs) and/or a combination of psychological therapy with a 'TCA'. The 'TCA' arm was included in the TCA class consisting of the individual TCA drugs that were part of the decision problem (i.e. amitriptyline and lofepramine). The combined psychological intervention plus 'TCA' was included in the respective combination class of the psychological intervention plus antidepressant.
2. Studies that included arms of a 'SSRI' (comprising a mixture of more than one SSRIs) and/or a combination of psychological therapy with a 'SSRI'. The 'SSRI' arm was included in the SSRI class consisting of the individual SSRI drugs that were part of the decision problem (i.e. citalopram, escitalopram, fluoxetine and sertraline). The combined psychological intervention plus 'SSRI' was included in the respective combination class of the psychological intervention plus antidepressant.
3. Studies that included arms of 'antidepressants' (comprising a mixture of more than one defined or undefined antidepressants) and/or a combination of psychological therapy with 'antidepressants'. The 'antidepressant' arm formed a separate node in the network, of no interest for the decision problem. However, it was decided to be retained as a separate node in the network as it provided links between psychological and combination interventions (and possibly other links between the interventions that had been compared with an 'antidepressant'). The psychological therapy plus 'antidepressant' combined intervention was classified under the respective combination class of the psychological therapy plus antidepressant.
4. Studies that assessed a combination of psychological therapy with a drug that was not considered in the NMAs versus psychological therapy alone or versus the specific drug alone or versus another intervention, active or inactive, that was considered in the NMAs. Any specific drug arm extracted from such studies was classified under the 'SSRI' class if it was an SSRI; the 'TCA' class if it was a TCA; and the 'antidepressant' class if it was neither a SSRI nor a TCA. The combination arms was classified under the respective combination class of the psychological intervention plus antidepressant.

The NMAs undertaken to address the 2 review questions covered in this chapter (i.e. interventions for people with less severe depression and interventions for people with more severe depression) included 351 studies comparing 81 pharmacological, psychological and physical interventions alone or in combination.

Update 2017

7.3.2.13 Identifying antidepressants for inclusion in the NMAs

Given the potential size and complexity of the network, the GC agreed to focus on those antidepressants which were most likely to be considered for use as first-line interventions in the English healthcare system. In doing so the GC drew on a number of principles to guide their choice of specific antidepressants. These principles included:

- the existing evidence of differential efficacy of antidepressants from existing NMAs (e.g. Cipriani et al. 2009; Khoo et al. 2015)
- the existing evidence on the tolerability of different antidepressants (e.g. Cipriani et al. 2009; Khoo et al. 2015)
- safety, including toxicity in overdose (e.g. Buckley and McManus 2002)
- other effects of antidepressants including sedative properties, discontinuation problems. weight gain or interactions with other drugs (e.g. Watanabe et al. 2011)
- the requirement to have a range of different drugs available for individuals who cannot tolerate a particular drug or where previous experience indicates a particular antidepressant or class of antidepressants are more or less effective.

In addition, the GC took a number of other factors into consideration including current usage of antidepressant drugs using data on current levels of prescribing, which indicated that citalopram, fluoxetine, sertraline, mirtazapine and amitriptyline were, in rank order, the 5

most commonly prescribed antidepressants (based on CPRD [Clinical Practice Research Datalink] antidepressant usage data provided by the GC, referring to 7,272 people with a first-ever episode of depression presenting to 141 practices in England between April 2011 and May 2012; usage of antidepressants prescribed for other conditions [such as pain, insomnia, migraine, etc.] were excluded from this dataset; patients' level of depressive symptom severity was not reported in the dataset). These drugs were reviewed against the principles set out above and it was decided to include them all. In addition, imipramine was included, not as a possible first-line treatment but because its use as a comparator in a large number of drug trials meant that it served to strengthen the links in the network. Other drugs were considered but were excluded from both NMAs (i.e. for people with less severe depression and for people with more severe depression), for example venlafaxine and paroxetine on the grounds of discontinuation symptoms (Schatzberg et al. 2006); agomelatine because of the additional monitoring requirements; reboxetine because of concerns about its efficacy; duloxetine, fluvoxamine and trazodone because of their limited current use and vortioxetine as it is recommended by NICE as a third-line agent. The majority of the TCAs (with the exception of amitriptyline, which was among the top-5 most commonly prescribed antidepressants) were excluded on the grounds of increased toxicity with the exception of lofepramine which was included on the grounds of the evidence of less toxicity in overdose. Nortriptyline was not included in the network but was assessed in a separate pairwise meta-analysis because the GC were interested in its potential use in older people with depression.

7.3.3 Class models

The NMAs that informed the review questions covered in this chapter utilised class models; this approach had two benefits: a. strength could be borrowed across interventions in the same class b. networks that were otherwise disconnected were possible to connect via interventions belonging to the same class. With the exception of one outcome (remission in responders in the less severe population), in all other cases random class effect models were used which assume that the effects of interventions in a class are distributed around a common class mean with a within-class variance. Under this approach individual treatment effects are drawn towards a class mean but individual intervention estimates are more precise.

Update 2017

Depending on the outcome assessed and the availability of respective data, classes were formed by a different number of interventions, ranging from one to eight. For interventions belonging to classes consisting of more than two interventions the pooled relative treatment effects were assumed to be exchangeable within class, with vague priors given to within-class mean treatment effects and informative priors given to within-class variability. For interventions belonging to a class formed only of themselves or one more intervention, the relative treatment effects were assumed to come from a normal distribution defined by the within-class mean treatment effects and variance being borrowed from another similar class in the model, where possible. Exceptions to this rule were interventions assumed to comprise their own single-intervention class, such as mirtazapine, pill placebo, attention placebo and wait list. In such cases, the mean and variance of the class were the same as the mean and variance of the intervention. For other classes consisting of one or two interventions, the assumptions for borrowing variance from similar classes were based on GC expert opinion. Details on the estimation of the variability within class and the assumptions used for classes borrowing variance from other classes are provided in Chapter 17, Section 17.2.3.

7.3.4 Data extracted, NMA outcomes and methods of outcome synthesis

For each RCT included in the NMA the following outcomes were extracted from each arm to inform NMAs on one or more outcomes:

- Number of participants randomised
- Numbers of participants discontinuing treatment (not completing the study)

- Number of participants discontinuing treatment due to the development of side effects
- Number of people responding to treatment, according to a minimum % change in score from baseline on a depressive symptom scale; in the majority of studies response was defined as a 50% reduction in score from baseline.
- Number of people remitting, defined as achieving a score below a pre-defined cut-off point on a depressive symptom scale.
- Mean change in score on a depressive symptom scale (and standard deviation or standard error of change score) from baseline; alternatively, mean baseline and endpoint continuous scale score data (and standard deviation or standard error of the scores) if change scores were not available. Relevant data were extracted for those randomised (intension-to-treat analysis, ITT) or study completers or both, as available.

Dichotomous and continuous data were extracted if they referred to a range of depressive symptom scales selected by the GC. Only data from one scale were extracted. If one RCT reported dichotomous or continuous data on more than one of the selected depression scales, then a hierarchy of depression scales was considered, and available data from the depression scale that was at a higher place in this hierarchy were extracted. The following depression scales (in the following hierarchy) were considered in the NMA, based on GC expert advice:

- MADRS
- HAMD
- QIDS
- PHQ-9
- CGI-I (Clinical Global Impressions – Improvement scale)
- CES-D
- BDI-I or BDI-II
- HADS-D
- HADS

CGI-I data were considered only in relation to the dichotomous outcome of response, which was defined as much or very much improved. Continuous data based on the CGI-I or the CGI-S (Clinical Global Impressions – Severity scale) were not extracted.

For each review question, a number of different outcomes were synthesised using NMA, which informed either the clinical or the economic analysis. For the clinical analysis, outcomes in those randomised based on an ITT approach were preferred. In contrast, the economic analysis required information on the conditional probability of outcomes (i.e. probability of outcomes based on the occurrence of a previous outcome, such as discontinuation or treatment completion) so that the sum of people across all model branches equalled the initial hypothetical cohort receiving each intervention of interest.

The following efficacy outcomes were considered for the clinical analysis:

- Standardised mean difference of depressive symptom scores (SMD); this outcome was used to combine evidence from studies reporting efficacy in terms of a continuous measurement on various depression scales, and was selected as the main clinical outcome by the GC. It was not used in the economic analysis
- Response in those randomised; this was selected as a secondary efficacy outcome
- Remission in those randomised; this was selected as a secondary efficacy outcome

The following conditional outcomes were selected to mainly inform the economic analysis:

- Treatment discontinuation for any reason in those randomised
- Treatment discontinuation due to side effects in those who discontinued treatment

Update 2017

- Response in those who completed treatment
- Remission in those who responded to treatment; this conditional outcome was originally chosen to inform the economic analysis. However, as the respective network in adults with a new episode of more severe depression was disconnected, the economic analysis was not informed by this outcome, and the outcome of remission in those who completed treatment was used instead. More details are provided in Chapter 14, where economic modelling methods and results are reported.

For the estimation of SMD of depressive symptom scores, the following extracted data were utilised, in the following hierarchy, depending on what was available in each study, in order to maximise the available information:

- mean change from baseline (CFB), standard deviation in CFB and total number of individuals in each arm (or the standard error of the mean change from baseline).
- baseline and endpoint mean scores, standard deviations and number of individuals, for each arm
- number of individuals responding to treatment in each arm, out of the total number of individuals

Details on data synthesis in order to obtain the SMD outcome are reported in Chapter 17, Section 17.2.5. Further information on the methods for estimation of within-study correlation and standard deviation at follow-up, which were essential for the estimation of the SMD outcome, is provided in Section 17.2.7.

For the estimation of response (either in those randomised or in completers), the following extracted data were utilised, in the following hierarchy, depending on what was available in each study, in order to maximise the available information:

- number of individuals responding to treatment in each arm, out of the total number of individuals in the arm
- mean CFB, standard deviation in CFB and total number of individuals in each arm (or the standard error of the mean change from baseline); estimated SMDs from these data were converted into Log-Odds Ratios (LORs) of response
- baseline and endpoint mean scores, standard deviations and number of individuals, for each arm; estimated SMDs from these data were converted to LORs of response.

Details on data synthesis in order to obtain the response outcome are provided in Chapter 17, Section 17.2.6.

For the estimation of remission (either in those randomised or in responders) only dichotomous remission data were utilised, due to disagreement between continuous and dichotomous remission data; details on this issue are provided in Chapter 17, Section 17.2.8.

It needs to be noted that in studies that reported change scores or endpoint continuous data for people randomised, some method of imputation of missing data for people who discontinued the study had been used, such as last observation carried forward (LOCF), baseline observation carried forward (BOCF), multiple imputation, etc. There is considerable variability in the underlying assumptions characterising each method of imputation; for example, LOCF and multiple imputation use different assumptions from BOCF; the latter corresponds to the assumption used to estimate dichotomous response in those randomised, i.e. that study non-completers do not respond (since they are counted as non-responders).

Data reported for the whole study sample (ITT analysis) that were based in LOCF were prioritised for extraction over data estimated based on other imputation methods, when more than one imputation methods were used in the study, but in general the extracted data reflected the available method of imputation in each RCT. This mixture of methods of imputation of missing continuous data may have potentially biased the outputs of the NMAs that utilised continuous data, i.e. the analyses reporting SMD of depressive symptom scores

Update 2017

and response in those randomised in both study populations, and this limitation needs to be taken into account when interpreting the outputs of these analyses. In contrast, the response in completers analyses do not suffer from this limitation, because the continuous data utilised in these NMAs were derived from study completers, so imputation of missing data was not required. Similarly, remission analyses (in those randomised, in completers and in responders) have only utilised dichotomous remission data, so this limitation is not relevant to them.

The studies and data that were used in the NMAs for every outcome of interest are provided in Appendix T.

7.3.5 Estimation, assessment of goodness of fit and inconsistency checks

Model parameters were estimated within a Bayesian framework using Markov chain Monte Carlo simulation methods implemented in WinBUGS 1.4.3 (Lunn et al., 2000; Spiegelhalter, 2001). In order to test whether prior estimates had an impact on the results, two chains with different initial values were run simultaneously. Convergence was assessed by inspection of the Brooks-Gelman–Rubin diagnostic plot and was satisfactory by 60,000 simulations for all outcomes. A further simulation sample of at least 40,000 iterations post-convergence was obtained on which all reported results were based.

Goodness of fit was tested using the posterior mean of the residual deviance, which was compared with the number of data points in the model. The Deviance Information Criterion (DIC) was also checked (Dias et al. 2011).

The between studies standard deviation (heterogeneity parameter) was estimated to assess the degree of statistical heterogeneity.

Consistency between the different sources of indirect and direct evidence was explored statistically by comparing the fit of a model assuming consistency with a model which allowed for inconsistency (also known as an unrelated treatment effect model). The latter is equivalent to having separate, unrelated meta-analyses for every pair-wise contrast but assumes a common between-study heterogeneity across all comparisons. The inconsistency model did not assume any class relation between interventions. If the inconsistency model had the smallest posterior mean residual deviance or heterogeneity then this indicated potential inconsistency in the data.

Details on the methods of testing for goodness of fit are reported in Chapter 17, Section 17.2.4.

7.3.6 Bias adjustment models

Publication bias is known to affect results of meta-analyses in several clinical areas, including Depression (Driessen et al., 2015; Moreno et al., 2009 & 2011; Trinquart et al., 2012; Turner et al., 2008). Small size studies are associated with publication bias (small studies with positive results are more likely to be published compared with small studies with negative results) and may also be associated with lower study quality. It has been shown that published smaller studies tend to overestimate the relative treatment effect of interventions vs control, compared to larger studies (Chaimani et al., 2013; Moreno et al., 2011). Regression using a measure of study precision has been successfully employed in published literature to adjust for small study effects in meta-analysis, with the study variance of the treatment effect, which is a measure of the latter's precision, being typically used to adjust for study size (Chaimani et al., 2013; Moreno et al., 2011).

As the NMAs included a significant number of small studies, sensitivity analyses were carried out on selected outcomes, which adjusted for bias associated with small study size effects. The analyses, which were based on the assumption that the smaller the study the greater the bias, attempted to estimate the "true" treatment effect, which would be obtained in a study of

Update 2017

infinite size. This was taken to be the intercept in a regression of the treatment effect against the study variance. The GC expressed the opinion that bias would act to favour active interventions when compared with an inactive control, but that there would be no systematic preference for comparisons between active interventions. These assumptions were supported by empirical evidence of the direction and magnitude of small study bias in meta-analyses of psychological interventions versus control (Driessen et al., 2015) and of anti-depressants versus placebo (Turner et al., 2008).

Bias adjustment models were therefore developed to estimate a potentially non-zero mean bias, with an estimated variance, for comparisons of active interventions to controls, while forcing the mean bias to be zero in active versus active comparisons, whilst still allowing a non-zero variance around this zero mean. This was to allow for the fact that small studies may exaggerate effects of one active intervention over another, but that this exaggerated effect may cancel out across multiple studies, with no particular intervention being favoured over another across all studies.

Bias adjustment models were applied to both populations (adults with less severe and adults with more severe depression) onto the following outcomes synthesised in NMAs:

- SMD of depressive symptom scores
- Treatment discontinuation for any reason in those randomised
- Response in completers

SMD of depressive symptom scores was selected for sensitivity analysis as it was the main efficacy outcome considered by the GC. The other two outcomes were selected for sensitivity analysis because they were the main NMA outcomes that informed the economic analysis, with the highest anticipated impact on the results. Subsequently, a probabilistic sensitivity analysis was conducted using the outputs of the bias-adjusted NMAs on these two outcomes, as reported in Chapter 14, section 14.2.12 (methods) and 14.3 (results).

Full details on the methods used to develop and test bias NMA models are reported in Appendix N.

Update 2017

7.3.7 Presentation of the results – selection of baseline comparator (reference)

Results of the NMAs are reported as posterior mean SMD of depressive symptom scores or LORs (for dichotomous data), as appropriate, with 95% Credible Intervals (CrI) compared with pill placebo, which was the baseline selected by the GC, as it is well-defined across trials and has its own established effect. In contrast, the definition of treatment as usual may vary from crisis intervention through a regular antidepressant treatment to a GP visit when needed, and was therefore deemed a sub-optimal baseline comparator. Wait list was considered to have a minimal effect and to potentially hinder other underlying interventions within the wait list arms across studies and therefore was also deemed an inappropriate baseline comparator. The GC considered the comparisons of psychological interventions and classes with pill placebo as an advantage of conducting the NMAs, because psychological therapies are not routinely compared with pill placebo, unless active drug arms are included in the trial. A further advantage of selecting pill placebo is that it provides a more conservative estimate and convincing comparison for clinical effect and addresses treatment expectancy effects for interventions.

This chapter provides a summary of the NMA results on outcomes considered for the clinical analysis. The networks, numbers randomised and relative effects versus pill placebo are reported for classes of interventions for all outcomes informing the clinical analysis; they are also illustrated in forest plots. In addition, posterior mean ranks of each class (and 95% CrI) are provided, in which lower rankings suggest a better outcome. Only classes of interest (i.e. being part of the decision problem) were included in the calculations of the rankings. For SMD of depressive symptom scores, which was the main efficacy outcome, the forest plots of individual intervention effects versus pill placebo are also provided for information.

Furthermore, the relative effects versus TAU are provided on the SMD outcome for comparison with relative effects versus pill placebo.

An overview of the results on outcomes used in the economic analysis (in terms of posterior mean odds ratios and 95% CrI of interventions of interest versus pill placebo) are reported in the respective economic modelling chapter (Chapter 14, section 14.2.5).

Detailed results of the NMAs on all outcomes that informed the clinical and the economic analysis are reported in Chapter 17.

7.3.8 Subgroup analyses

Sufficient data were available to conduct sub-analyses of RCTs conducted in inpatient versus outpatient populations, and older (>60 years of age) versus younger (<60 years of age) adults. Data for these sub-analyses were pooled across review questions 2.1 and 2.2 to allow for comparison of differential effects in different populations, thereby more helpfully informing GC decision making. The results of these analyses are provided below.

7.4 Review question

- For adults with a new episode of less severe depression, what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination for the treatment of depression?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 43. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Update 2017

Table 43: Clinical review protocol summary for the review of acute treatment for less severe depression

Component	Description
Review question	For adults with a new episode of less severe depression, what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination for the treatment of depression? (RQ2.1)
Population	• Adults receiving first line treatment for a new episode of depression, as defined by a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on scales (and including those with subthreshold depressive symptoms). If some, but not all, of a study's participants are eligible for the review, for instance, mixed anxiety and depression diagnoses, and we are unable to obtain the appropriate disaggregated data, then we will include a study if at least 80% of its participants are eligible for this review Baseline mean scores are used to classify study population severity according to less severe (RQ 2.1) or more severe (RQ 2.2) using the thresholds outlined in Table 42. If baseline mean scores are not available, severity will be classified according to the inclusion criteria of the study or the description given by the study authors (but only in cases where this is unambiguous, i.e. 'severe' or 'subthreshold' or 'mild').
Intervention(s)	The following interventions will be included in the NMA: **Psychological interventions**:

Component	Description
	• Behavioural therapies (including behavioural activation, behavioural therapy [Lewinsohn 1976], coping with depression course [individual and group] and social rhythm therapy [SRT]) • Cognitive and cognitive behavioural therapies (including CBT individual or group [defined as under or over 15 sessions], problem solving, rational emotive behaviour therapy [REBT] and third-wave cognitive therapies individual or group) • Counselling (including directive counselling, emotion-focused therapy [EFT], non-directive counselling and relational client-centred therapy) • Interpersonal psychotherapy • Psychodynamic psychotherapies (including individual or group-based short-term psychodynamic psychotherapy, long-term psychodynamic psychotherapy and psychodynamic counselling) • Psychoeducational interventions (including psychoeducational group programmes, intensive clinical management and lifestyle factors discussion) • Self-help with or without support (including cognitive bibliotherapy with or without support, computerised CBT [CCBT] with or without support, computerised problem solving therapy with or without support, computerised psychodynamic therapy with or without support, online positive psychological intervention and self-examination therapy) **Pharmacological interventions:** • SSRIs (citalopram, escitalopram, sertraline, fluoxetine) • TCAs (amitriptyline, lofepramine) • Other antidepressant drugs (mirtazapine) Note that in order to maximise connectivity in the network specific drugs that are excluded and 'any antidepressant' or 'any SSRI' or 'any TCA' nodes will be added where they have been compared against a psychological intervention and/or combined with a psychological intervention but they will not be considered as part of the decision problem. **Physical interventions:** •Exercise (including yoga) The following interventions may be compared in pairwise comparisons (however will not be included in the NMA): • Acupuncture • Behavioural couples therapy • Light therapy (for depression but not for SAD) • Nortriptyline (for older adults) • Omega-3 fatty acids • Psychosocial interventions (including befriending, mentoring, peer support and community navigators)
Comparison	• Any other intervention • Treatment as usual • Waitlist • Placebo • Imipramine
Critical outcomes	Critical outcomes **Efficacy:**

Update 2017

Component	Description
	• Depression symptomology (mean endpoint score or change in depression score from baseline) • Remission (usually defined as a cut off on a depression scale) • Response (e.g. reduction of at least 50% from the baseline score on HAMD/MADRS) **Acceptability/tolerability**: • Discontinuation due to side effects (for pharmacological trials) • Discontinuation due to any reason (including side effects) • The following depression scales will be included in the following hierarchy: i. MADRS ii. HAMD iii. QIDS iv. PHQ v. CGI vi. CES-D vii. BDI viii. HADS-D (depression subscale) ix. HADS (full scale) Only one continuous scale will be used per study • For studies reporting response and/or remission, the scale used in the study to define cut-offs for response and/or remission will be used • If more than one definition is used, a hierarchy of scales will be adopted (hierarchy listed above) • For studies not reporting dichotomous data, a hierarchy of scales will be adopted for continuous outcomes
Study design	• Systematic reviews of RCTs • RCTs • Cluster RCTs

Update 2017

7.4.1 Clinical evidence

7.4.1.1 Study characteristics

1372 studies were considered at full text for inclusion in this review. Of these, 205 RCTs (k=205, n=28,047) were included in this network meta-analysis.

Of the 205 RCTs included within this network and reporting either a HAM-D or MADRS score at baseline, the mean depression severity scores were HAM-D=19.7 (n=75) and MADRS=22.6 (n=17) respectively. 24 were UK based RCTs.

For a full list of included and excluded studies, study characteristics of included studies and risk of bias please see Appendix J3.1 and J3.2.

Data were not available for every outcome of interest for the majority of included RCTs. For the outcomes considered in the clinical analysis, the following information was available:

- SMD of depressive symptom scores: 20 trials reported CFB data; 76 trials reported mean baseline and endpoint symptom scores and another 10 reported dichotomous response data. In total, 106 RCTs provided data on 15,671 trial participants that were used to inform the SMD outcome.

- Response in those randomised: 52 studies reported dichotomous response data, another 10 reported CFB data and in 68 studies baseline and endpoint symptom scores were available. In total, 130 RCTs with data on 19,320 participants informed this outcome
- Remission in those randomised: 65 studies provided dichotomous remission data on 10,179 participants
- Relevant information on the number of studies and study participants that provided data on the outcomes that were used to inform the economic analysis are provided in Chapter 17, in respective outcome sections. The studies and data that were used in the NMAs for every outcome of interest are provided in Appendix T.

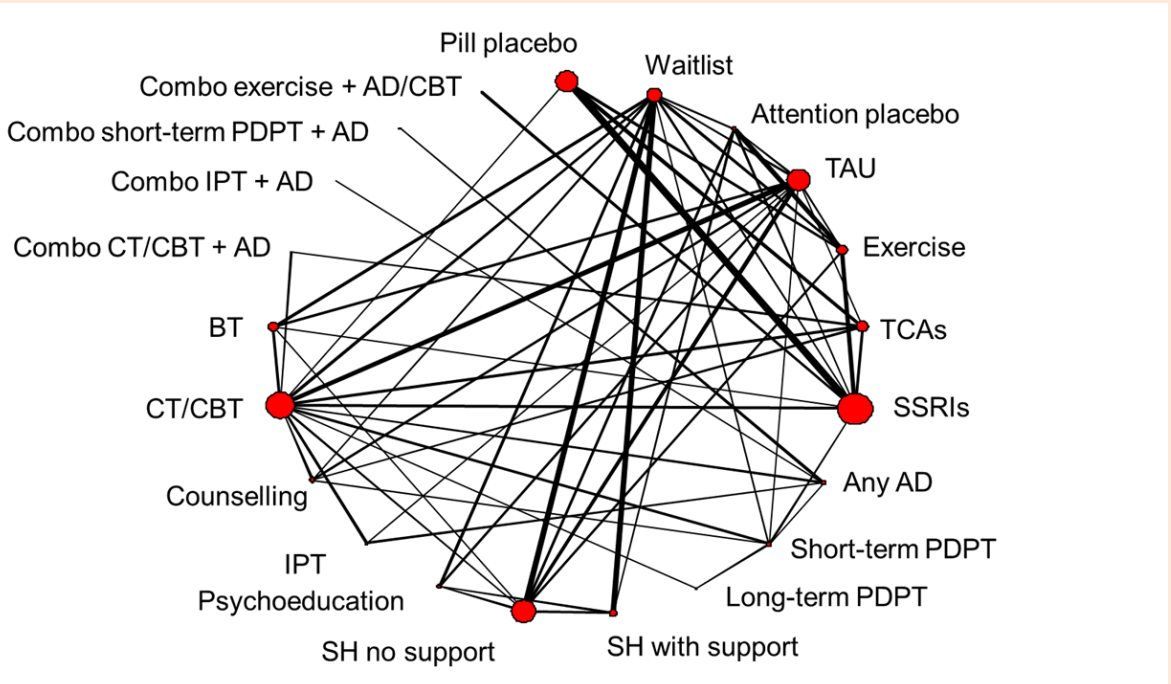
7.4.1.2 Results of the network meta-analysis

This section reports only NMA results that informed clinical evidence. Detailed NMA findings on all outcomes, including those that informed the economic analysis, are reported in the respective sections of Chapter 17.

Standardised mean difference (SMD) of depressive symptom scores

The network diagram of all studies included in this analysis by class is provided in Figure 3. The network diagram of the studies included in this analysis by intervention is provided in Chapter 17, Section 17.3.1.7.The relative effects of all classes versus pill placebo and versus TAU (posterior mean SMD with 95% CrI) are provided in Table 44, together with posterior mean ranks of each class (with 95% CrI). Classes in the table have been ranked from smallest to largest ranking (with lower rankings suggesting better outcome). The relative effects of every class versus pill placebo and of every intervention versus pill placebo are shown in Figure 4 and Figure 5, respectively. Detailed results are provided in Chapter 17, Sections 17.3.1.7, 17.9 and 17.10.

Figure 3 Network diagram of all studies included in the analysis of standardised mean difference (SMD) of depressive symptom scores in people with a new episode of less severe depression by class



Update 2017

Table 44 Results of NMA in people with a new episode of less severe depression. Standardised mean difference of depressive symptom scores: Posterior effects (SMD of depressive symptom scores) of all classes versus pill placebo and TAU and ranking of classes

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Effect vs TAU (mean, 95% CrI)	Mean Rank (95% CrI)
Combined (IPT + AD)	63	-1.48 (-2.86 to -0.10)	-1.67 (-3.03 to -0.30)	2.71 (1 to 13)
Combined (Short-term PDPT + AD)	165	-1.14 (-2.50 to 0.21)	-1.33 (-2.68 to 0.02)	4.23 (1 to 17)
Self-help with support	514	-0.77 (-1.52 to -0.03)	-0.96 (-1.69 to -0.24)	5.48 (1 to 14)
Long-term PDPT	128	-0.75 (-1.70 to 0.19)	-0.93 (-1.86 to -0.01)	5.99 (1 to 17)
Combined (Exercise + AD/CBT)	79	-0.59 (-1.87 to 0.70)	-0.77 (-2.06 to 0.53)	8.11 (1 to 20)
CT/CBT	2026	-0.46 (-0.85 to -0.07)	-0.64 (-0.99 to -0.29)	8.16 (4 to 13)
Combined (CT/CBT + AD)	36	-0.53 (-1.88 to 0.82)	-0.71 (-2.05 to 0.63)	8.68 (1 to 20)
Behavioural therapies	717	-0.43 (-1.29 to 0.44)	-0.61 (-1.45 to 0.23)	9.11 (2 to 19)
Psychoeducational interventions	268	-0.35 (-1.22 to 0.52)	-0.53 (-1.38 to 0.31)	10.05 (2 to 20)
Exercise	752	-0.34 (-1.57 to 0.89)	-0.52 (-1.75 to 0.70)	10.37 (1 to 20)
IPT	234	-0.29 (-1.18 to 0.59)	-0.47 (-1.34 to 0.39)	10.81 (3 to 20)
SSRIs	2463	-0.23 (-0.74 to 0.28)	-0.41 (-0.97 to 0.14)	11.38 (4 to 19)
Short-term PDPT	379	-0.23 (-0.89 to 0.45)	-0.41 (-1.04 to 0.24)	11.51 (4 to 19)
TCAs	811	-0.19 (-0.76 to 0.38)	-0.38 (-0.96 to 0.21)	11.93 (5 to 19)
Self-help without support	1721	-0.19 (-0.89 to 0.53)	-0.37 (-1.05 to 0.32)	12.02 (4 to 20)
Counselling	406	-0.13 (-1.00 to 0.75)	-0.31 (-1.16 to 0.54)	12.71 (3 to 20)
Pill placebo	1564	reference	-0.18 (-0.45 to 0.07)	14.90 (11 to 18)
Attention placebo	294	0.02 (-0.29 to 0.32)	-0.17 (-0.45 to 0.11)	15.09 (10 to 19)
TAU	1675	0.18 (-0.07 to 0.45)	reference	17.48 (14 to 19)
Waitlist	1035	0.38 (0.11 to 0.66)	0.20 (0.00 to 0.40)	19.31 (17 to 20)

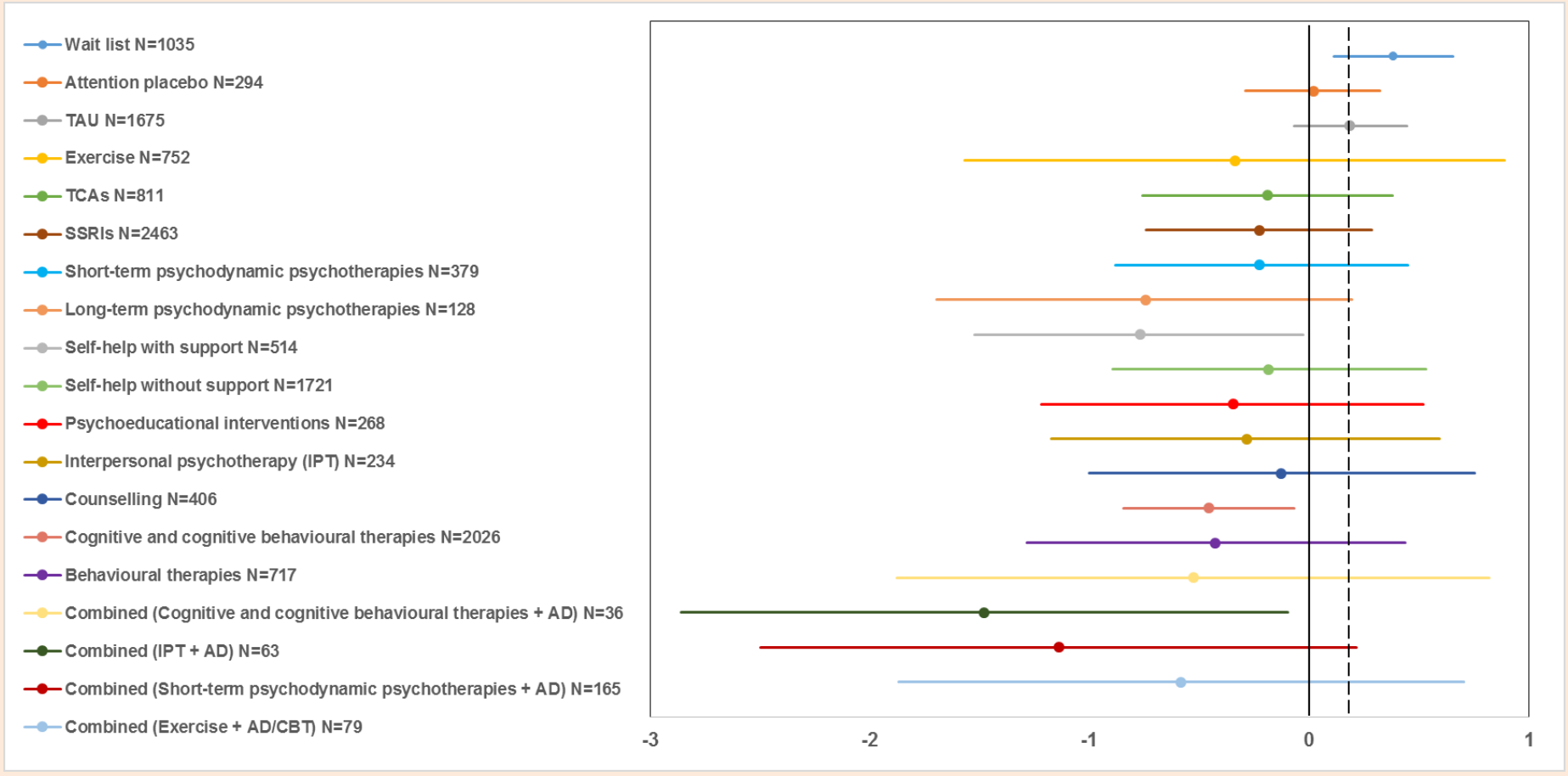
Notes:

Negative effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo or TAU)

AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants

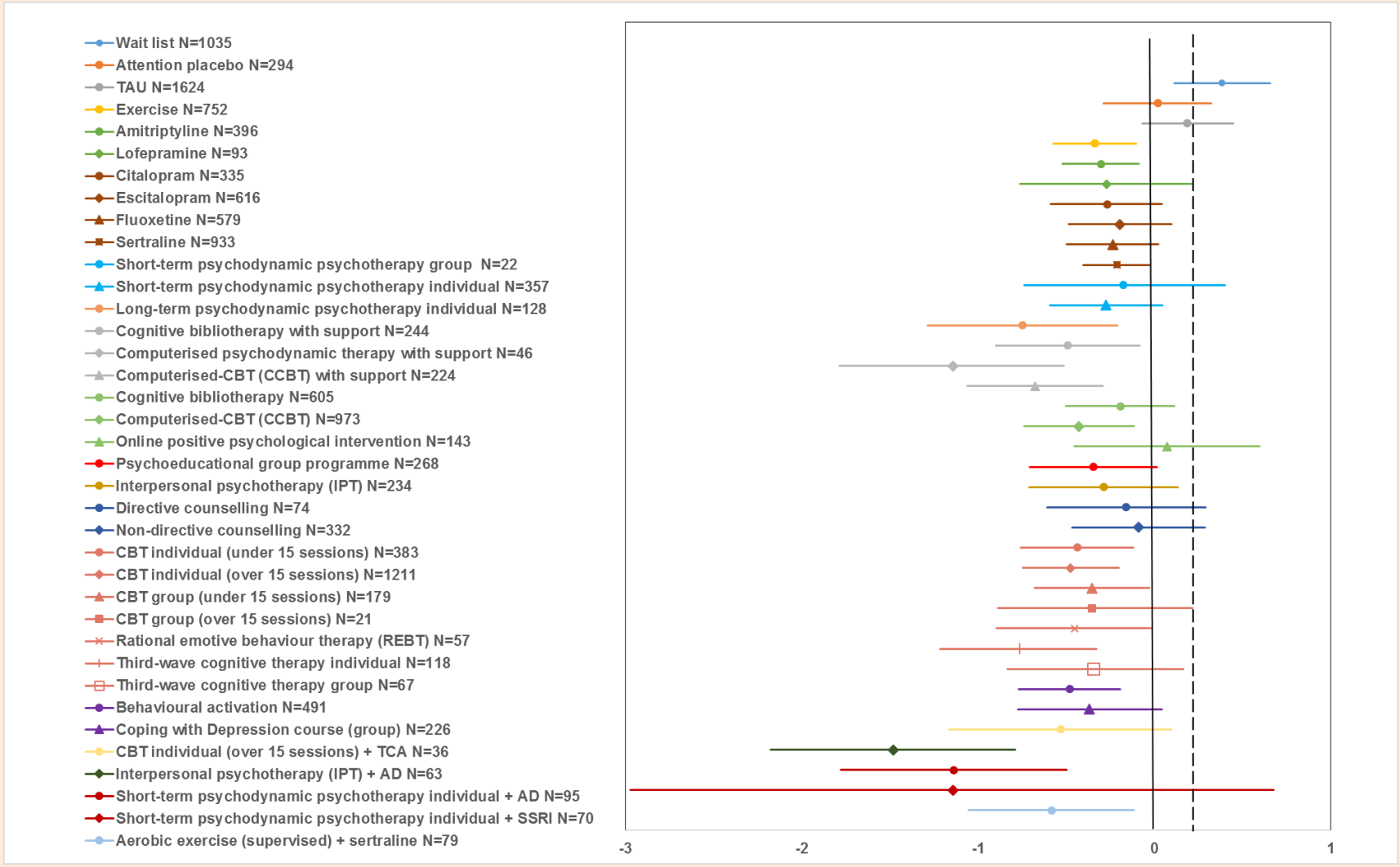
Update 2017

Figure 4 Results of NMA in people with a new episode of less severe depression. Standardised mean difference (SMD) of depressive symptom scores of all classes versus pill placebo (N=1564) [values on the left side of the vertical axis indicate a better effect compared with pill placebo; dotted line indicates TAU effect]



Update 2017

Figure 5 Results of NMA in people with a new episode of less severe depression. Standardised mean difference (SMD) of depressive symptom scores of all interventions versus pill placebo (N=1564) [values on the left side of the vertical axis indicate a better effect compared with pill placebo; dotted line indicates TAU effect]



Update 2017

Response in those randomised

The network diagram of all studies included in this analysis by class is provided in Figure 6. The network diagram of studies included in this analysis by intervention is provided in Chapter 17, Section 17.3.1.6.The relative effects of all classes versus pill placebo (posterior mean LORs with 95% CrI) are provided in Table 45, together with posterior mean ranks of each class (with 95% CrI). Classes in the table have been ranked from smallest to largest ranking (with lower rankings suggesting better outcome). The relative effects of every class versus pill placebo are shown in Figure 7. Detailed results are provided in Chapter 17, Section 17.3.1.6, 17.9 and 17.10.

Figure 6 Network diagram of all studies included in the analysis of response in those randomised in people with a new episode of less severe depression by class

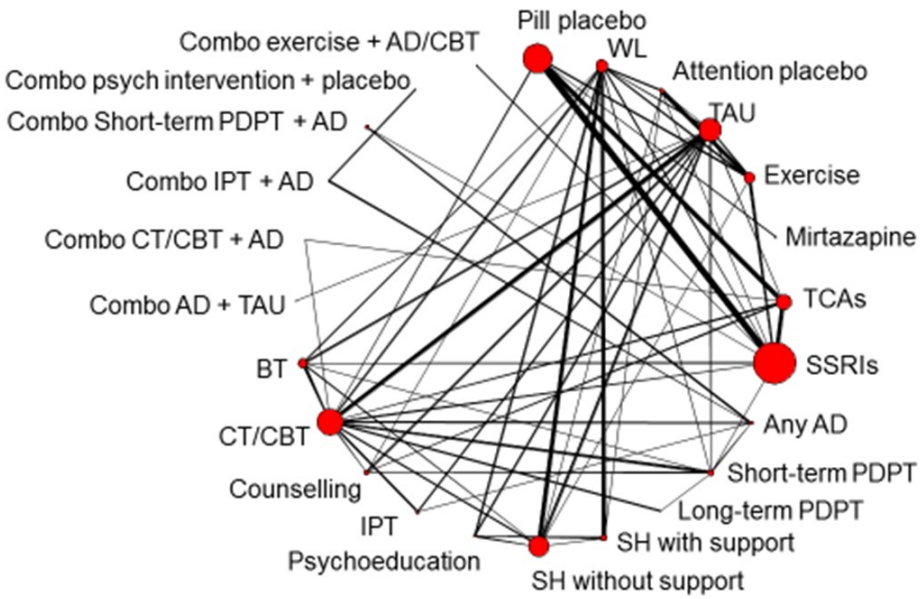


Table 45 Results of NMA in people with a new episode of less severe depression. Response in those randomised: Posterior effects (Log-Odds Ratios of response) of all classes versus pill placebo and ranking of classes

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Combined (IPT + AD)	76	2.51 (0.56 to 4.46)	2.73 (1 to 12)
Combined (Exercise + AD/CBT)	79	2.07 (0.34 to 3.81)	3.91 (1 to 15)
Combined (Short-term PDPT + AD)	295	1.79 (0.36 to 3.20)	4.55 (1 to 14)
Long-term PDPT	128	1.46 (-0.09 to 3.00)	6.29 (1 to 18)
Mirtazapine	45	1.36 (-0.20 to 2.99)	7.08 (1 to 19)
Combined (CT/CBT + AD)	36	1.36 (-0.44 to 3.16)	7.19 (1 to 19)
Behavioural therapies	764	1.14 (0.19 to 2.09)	7.60 (3 to 15)
CT/CBT	2147	0.98 (0.32 to 1.63)	8.56 (4 to 14)
Self-help with support	544	0.88 (-0.09 to 1.83)	9.67 (3 to 17)
Exercise	945	0.83 (-0.47 to 2.16)	10.25 (2 to 20)
Short-term PDPT	491	0.68 (-0.38 to 1.73)	11.43 (4 to 19)
SSRIs	3547	0.65 (0.05 to 1.26)	11.68 (6 to 17)

Update 2017

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Psychoeducational interventions	268	0.54 (-0.75 to 1.80)	12.52 (4 to 20)
IPT	234	0.54 (-0.74 to 1.85)	12.55 (4 to 20)
TCAs	1292	0.54 (-0.18 to 1.26)	12.70 (6 to 19)
Counselling	406	0.40 (-0.71 to 1.52)	13.74 (5 to 20)
Self-help without support	1736	0.32 (-0.60 to 1.22)	14.56 (7 to 20)
Attention placebo	381	0.14 (-0.51 to 0.79)	16.24 (11 to 20)
Pill placebo	2439	Reference	17.39 (14 to 20)
TAU	1949	-0.36 (-0.90 to 0.17)	19.51 (18 to 21)
Waitlist	1039	-0.76 (-1.38 to -0.17)	20.86 (20 to 21)

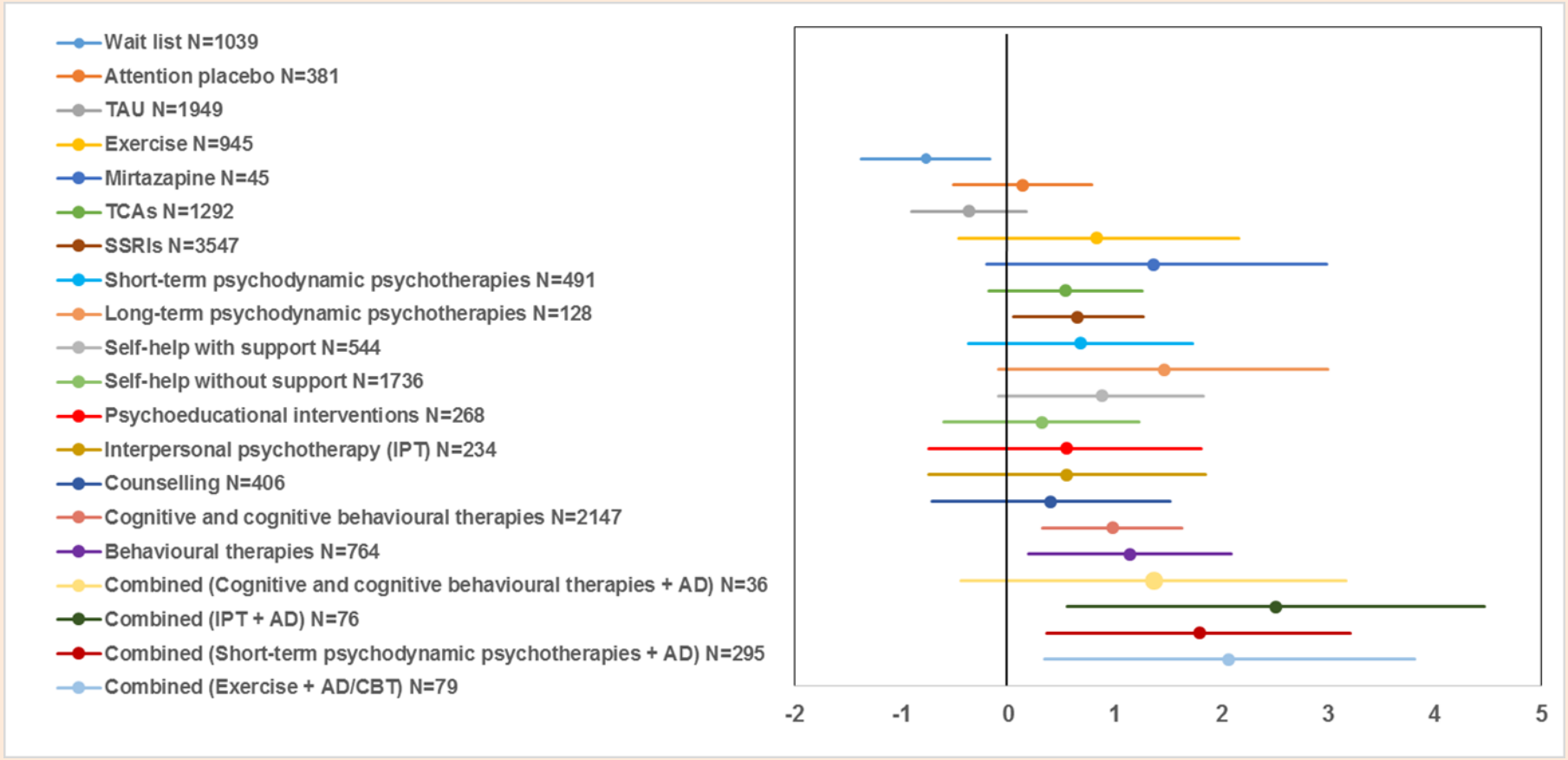
Notes:

Positive effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo)

AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants

Update 2017

Figure 7 Results of NMA in people with a new episode of less severe depression. Log-Odds Ratios of response in those randomised of all classes versus pill placebo (N=2439) [values on the right side of the vertical axis indicate a better effect compared with pill placebo]



Update 2017

Remission in those randomised

The network diagram of all studies included in this analysis by class is provided in Figure 8. The network diagram of the studies included in this analysis by intervention is provided in Chapter 17, section 17.3.1.4.The relative effects of all classes versus pill placebo (posterior mean LORs with 95% CrI) are provided in Table 46, together with posterior mean ranks of each class (with 95% CrI). Classes in the table have been ranked from smallest to largest ranking (with lower rankings suggesting better outcome). The relative effects of every class versus pill placebo are shown in Figure 9. Detailed results are provided in Chapter 17, Sections 17.3.1.4, 17.9 and 17.10.

Figure 8 Network diagram of all studies included in the analysis of remission in those randomised in people with a new episode of less severe depression by class

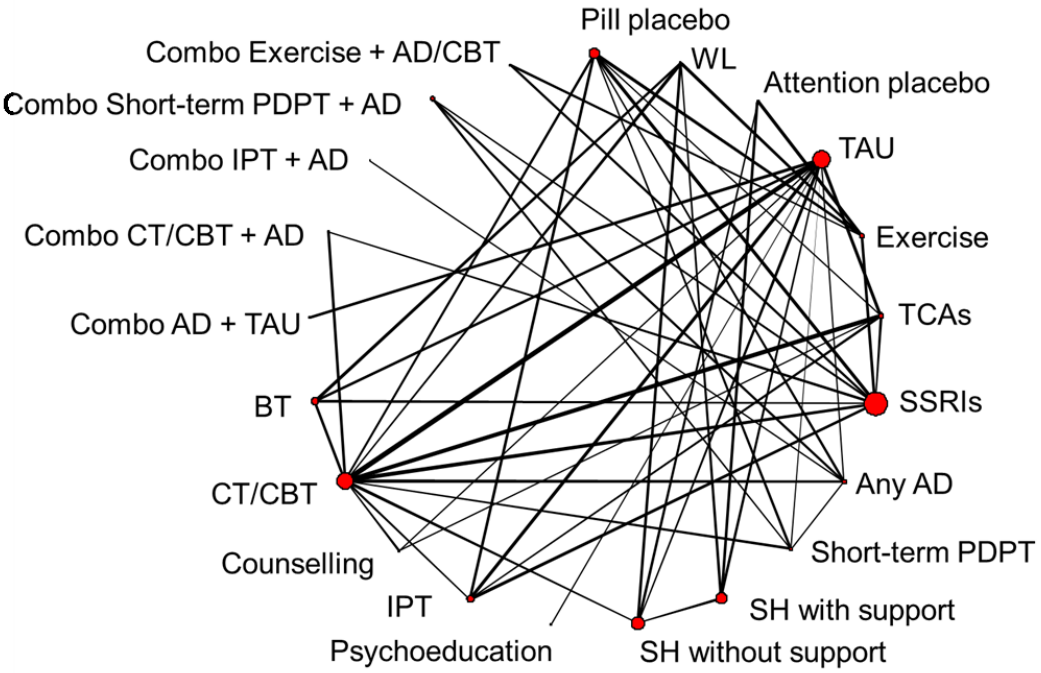


Table 46 Results of NMA in people with a new episode of less severe depression. Remission in those randomised: Posterior effects (Log-Odds Ratios of remission) of all classes versus pill placebo and ranking of classes

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Behavioural therapies	545	1.29 (0.39 to 2.23)	3.87 (1 to 9)
Counselling	64	1.47 (-0.33 to 3.28)	4.30 (1 to 16)
Combined (IPT + AD)	63	1.32 (-0.52 to 3.20)	5.06 (1 to 17)
Combined (Short-term PDPT + AD)	323	1.15 (-0.33 to 2.69)	5.49 (1 to 16)
CT/CBT	1196	0.82 (0.07 to 1.58)	6.88 (3 to 12)
IPT	519	0.85 (-0.37 to 2.07)	7.05 (1 to 16)
Psychoeducational interventions	119	0.87 (-0.79 to 2.56)	7.25 (1 to 17)
SSRIs	1683	0.60 (-0.09 to 1.28)	8.70 (3 to 14)
Combined (CT/CBT + AD)	84	0.55 (-0.65 to 1.76)	9.26 (2 to 17)
Self-help with support	832	0.54 (-0.64 to 1.72)	9.27 (2 to 17)

Update 2017

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Short-term PDPT	248	0.53 (-0.62 to 1.67)	9.39 (2 to 17)
TCAs	372	0.37 (-0.54 to 1.26)	10.68 (4 to 17)
Combined (Exercise + AD/CBT)	110	0.13 (-1.41 to 1.68)	12.06 (2 to 19)
Exercise	330	-0.01 (-1.42 to 1.40)	13.07 (3 to 19)
TAU	1258	0.03 (-0.63 to 0.71)	13.68 (10 to 17)
Self-help without support	926	-0.06 (-1.20 to 1.04)	13.81 (6 to 18)
Pill placebo		reference	13.86 (9 to 17)
Attention placebo	204	-0.89 (-1.87 to 0.08)	17.76 (15 to 19)
Waitlist		-1.23 (-2.23 to -0.23)	18.56 (17 to 19)

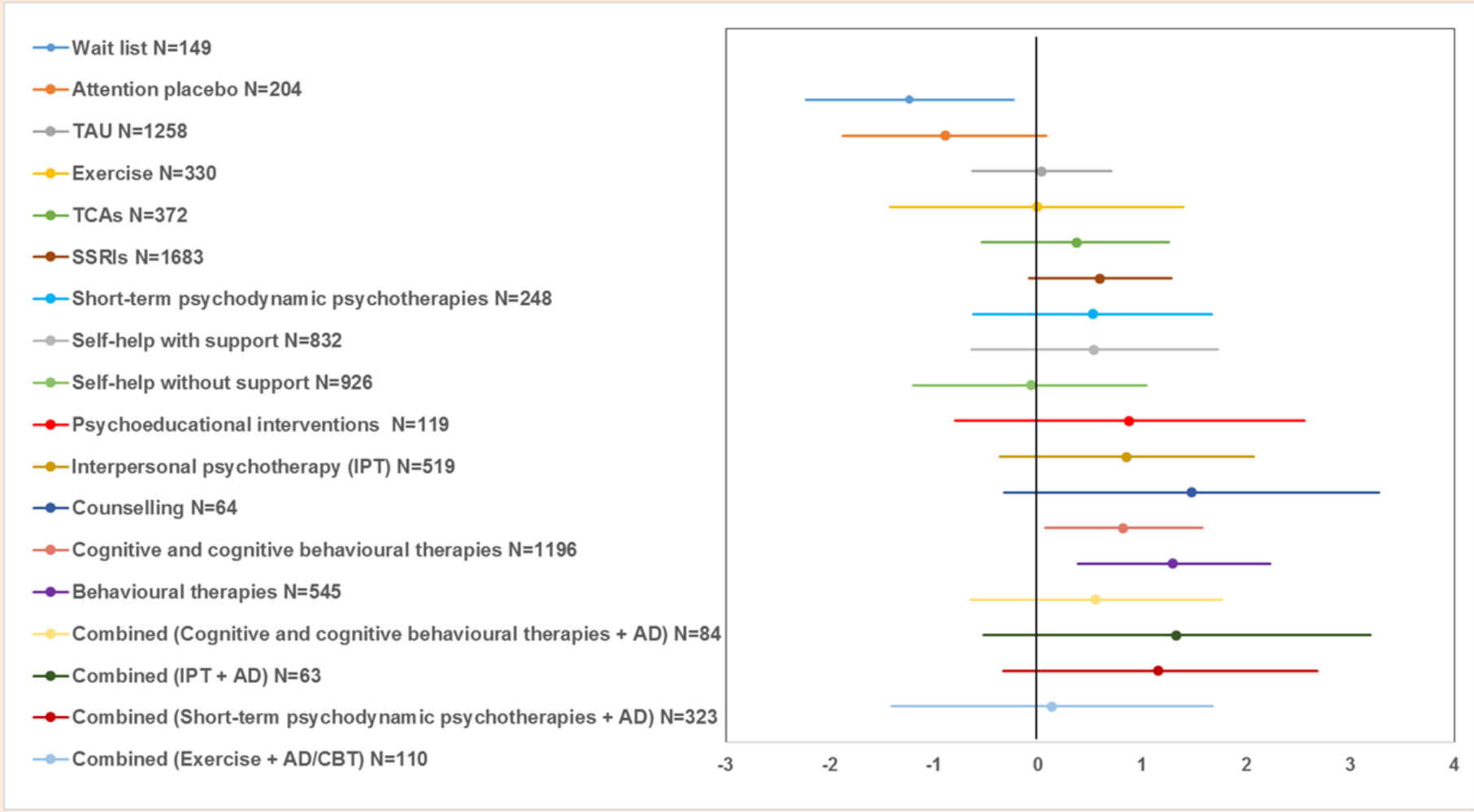
Notes:

Positive effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo)

AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants

Update 2017

Figure 9 Results of NMA in people with a new episode of less severe depression. Log-Odds Ratios of remission in those randomised of all classes versus pill placebo (N=719) [values on the right side of the vertical axis indicate a better effect compared with pill placebo]



Update 2017

A comparison of the results of the NMAs across the 3 outcomes of SMD of depressive symptom scores, response in those randomised and remission in those randomised can be made by inspection of Table 47. It can be seen that ranking of interventions and effects versus pill placebo were not consistent across analyses:

- Results for pharmacological classes of interventions (SSRIs and TCAs) were broadly consistent across the 3 analyses, although their ranking was somewhat better on the outcome of remission in those randomised; mirtazapine ranked highly on the response in those randomised outcome, which was the only outcome for which data on mirtazapine were available.
- Self-help without or with minimal support showed small or no benefit across the 3 analyses. Self-help with support showed a benefit across all 3 analyses; it ranked highly on the SMD, but was placed in lower rankings on the other two outcomes. Psychoeducation showed a small to moderate effect across all analysis, although its ranking ranged across the 3 analyses.
- Regarding high-intensity psychological classes of interventions, CT/CBT showed broadly consistent benefits across all analyses and had relatively high rankings (5-8). Behavioural therapies showed a benefit across all analyses and ranked first in the remission in those randomised analysis, but in lower places (7-8) in the other two analyses. Similarly, counselling ranked second best intervention in the remission in those randomised analysis, but showed small effects and had low rankings in the other two analyses. IPT showed also a higher effect (and ranking) in the remission in those randomised analysis compared with the other two analyses. Short-term PDPT showed a similar, low to medium effect across all 3 analyses. Long-term psychodynamic psychotherapy showed a large benefit and was ranked fourth in both the SMD and response in those randomised analyses; no remission data were available for long-term psychodynamic psychotherapy.
- Exercise showed a moderate effect and ranking in the SMD and response in those randomised analyses, and no effect (and, consequently, a low ranking) in the remission in those randomised analysis.
- Combined classes of interventions demonstrated, on balance, the highest effects and rankings. Combined interpersonal therapy with antidepressants and combined short-term psychodynamic psychotherapy with antidepressants were the only two classes that ranked in the top 4 places for all 3 outcomes. Combined CT/CBT with antidepressants was ranked in places 6-9 across the 3 analyses. Finally, combined exercise with CBT/antidepressants showed moderate to high effects in the SMD and response in those randomised analyses, but practically no benefit in the remission in those randomised analysis.

It needs to be noted that the 3 analyses were informed by different datasets, which may explain the discrepancies in relative effects and class rankings observed across the 3 outcomes. Nevertheless, the SMD and response in those randomised analyses may have potentially shared some study data, as in studies not reporting continuous data, dichotomous response data, if available, were used in the estimation of SMD and, conversely, in studies not reporting dichotomous response data, continuous symptom scale data, if available, were used in the estimation of response in those randomised. In contrast, the remission in those randomised analysis utilised different data from the other two analyses, which, in part, explains the considerable discrepancies observed in the results of some classes between this and the other two analyses.

Update 2017

1

Table 47 Comparison of NMA results across the outcomes considered in clinical analyses for people with a new episode of less severe depression: posterior effects of all classes versus pill placebo

Effect of every class versus pill placebo (mean, 95% CrI); classes listed according to their mean ranking (lowest to largest) for each outcome					
SMD of depressive symptom scores		**Response in those randomised (LORs)**		**Remission in those randomised (LORs)**	
Combined (IPT + AD)	-1.48 (-2.86 to -0.10)	Combined (IPT + AD)	2.51 (0.56 to 4.46)	Behavioural therapies	1.29 (0.39 to 2.23)
Combined (Short-term PDPT + AD)	-1.14 (-2.50 to 0.21)	Combined (Exercise + AD/CBT)	2.07 (0.34 to 3.81)	Counselling	1.47 (-0.33 to 3.28)
Self-help with support	-0.77 (-1.52 to -0.03)	Combined (Short-term PDPT + AD)	1.79 (0.36 to 3.20)	Combined (IPT + AD)	1.32 (-0.52 to 3.20)
Long-term PDPT	-0.75 (-1.70 to 0.19)	Long-term PDPT	1.46 (-0.09 to 3.00)	Combined (Short-term PDPT + AD)	1.15 (-0.33 to 2.69)
Combined (Exercise + AD/CBT)	-0.59 (-1.87 to 0.70)	Mirtazapine	1.36 (-0.20 to 2.99)	CT/CBT	0.82 (0.07 to 1.58)
CT/CBT	-0.46 (-0.85 to -0.07)	Combined (CT/CBT + AD)	1.36 (-0.44 to 3.16)	IPT	0.85 (-0.37 to 2.07)
Combined (CT/CBT + AD)	-0.53 (-1.88 to 0.82)	Behavioural therapies	1.14 (0.19 to 2.09)	Psychoeducation	0.87 (-0.79 to 2.56)
Behavioural therapies	-0.43 (-1.29 to 0.44)	CT/CBT	0.98 (0.32 to 1.63)	SSRIs	0.60 (-0.09 to 1.28)
Psychoeducation	-0.35 (-1.22 to 0.52)	Self-help with support	0.88 (-0.09 to 1.83)	Combined (CT/CBT + AD)	0.55 (-0.65 to 1.76)
Exercise	-0.34 (-1.57 to 0.89)	Exercise	0.83 (-0.47 to 2.16)	Self-help with support	0.54 (-0.64 to 1.72)
IPT	-0.29 (-1.18 to 0.59)	Short-term PDPT	0.68 (-0.38 to 1.73)	Short-term PDPT	0.53 (-0.62 to 1.67)
SSRIs	-0.23 (-0.74 to 0.28)	SSRIs	0.65 (0.05 to 1.26)	TCAs	0.37 (-0.54 to 1.26)
Short-term PDPT	-0.23 (-0.89 to 0.45)	Psychoeducation	0.54 (-0.75 to 1.80)	Combined (Exercise + AD/CBT)	0.13 (-1.41 to 1.68)
TCAs	-0.19 (-0.76 to 0.38)	IPT	0.54 (-0.74 to 1.85)	Exercise	-0.01 (-1.42 to 1.40)
Self-help without support	-0.19 (-0.89 to 0.53)	TCAs	0.54 (-0.18 to 1.26)	TAU	0.03 (-0.63 to 0.71)
Counselling	-0.13 (-1.00 to 0.75)	Counselling	0.40 (-0.71 to 1.52)	Self-help without support	-0.06 (-1.20 to 1.04)
Pill placebo	reference	Self-help without support	0.32 (-0.60 to 1.22)	Pill placebo	reference
Attention placebo	0.02 (-0.29 to 0.32)	Attention placebo	0.14 (-0.51 to 0.79)	Attention placebo	-0.89 (-1.87 to 0.08)

Effect of every class versus pill placebo (mean, 95% CrI); classes listed according to their mean ranking (lowest to largest) for each outcome					
SMD of depressive symptom scores		**Response in those randomised (LORs)**		**Remission in those randomised (LORs)**	
TAU	0.18 (-0.07 to 0.45)	Pill placebo	Reference	Waitlist	-1.23 (-2.23 to -0.23)
Waitlist	0.38 (0.11 to 0.66)	TAU	-0.36 (-0.90 to 0.17)		
		Waitlist	-0.76 (-1.38 to -0.17)		
Negative values favour classes on the left column		Positive values favour classes on the left column		Positive values favour classes on the left column	
AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; LORs: log-odds ratios; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants					

1

Update 2017

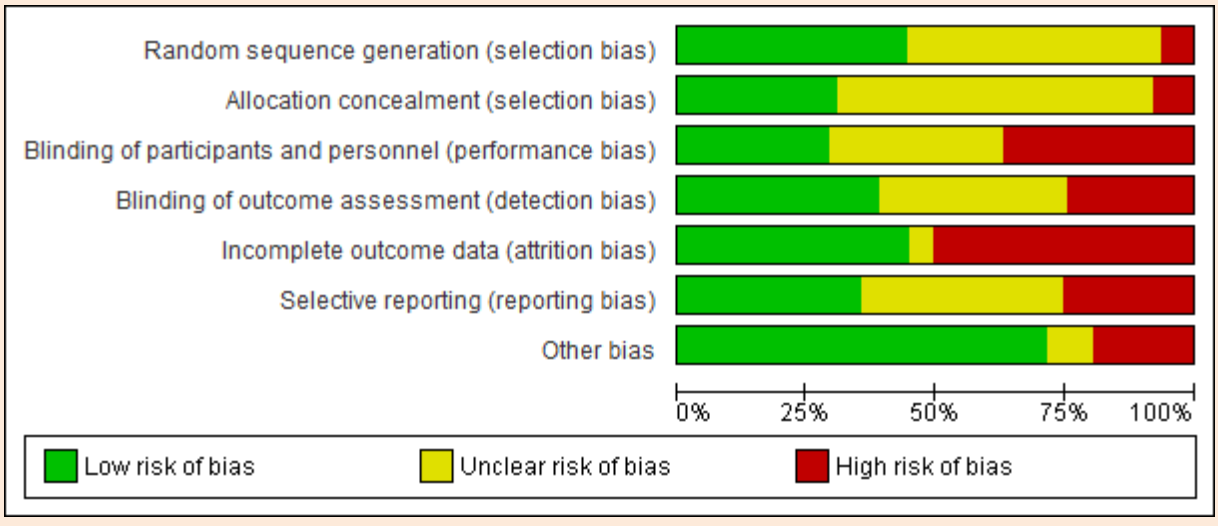
7.4.1.31 Quality of the evidence

The standard GRADE profiles for critical outcomes that have been used to rate the quality of evidence in pairwise meta-analyses conducted for this guideline have not been used for grading the quality in the NMA. This is because GRADE was not developed with network meta-analysis in mind and this is an area of methodological discussion and development. To evaluate the quality of the evidence of the NMAs undertaken to inform this guideline, we report information about the factors that would normally be included in a GRADE profile (i.e. risk of bias, publication bias, imprecision, inconsistency, and indirectness). Study quality and risk of bias were assessed for all studies, irrespective of whether they were included in the network meta-analysis or pairwise comparisons.

Risk of bias

We assessed all included trials for risk of bias (Appendix J3.2). Generally the standard of reporting in studies was quite low, as demonstrated by the risk of bias summary diagram below. Of the studies included in this NMA, 89 were at low risk for sequence generation and 60 of these were at low risk of bias for allocation concealment. Allocation concealment was unclear in 27 trials, and 2 trials were at high risk of bias. Trials of psychological therapies were typically considered at high risk of bias for participant and provider blinding per se; 59 trials were at low risk of bias for blinding participants and providers, although the rate of side effects may make it difficult to maintain blinding in pharmacological trials as well. Most reported outcomes were investigator-rated, and assessor blinding was considered separately for all trials; 78 at low risk of bias, 73 were unclear, and high risk in 49 trials. For incomplete outcome data, 92 trials were at low risk of bias; unclear risk in 8 trials, and 100 trials were at high risk of bias. Other sources of bias, potential or actual were identified in 56 RCTs. A summary is shown in Figure 10.

Figure 10: Risk of bias summary for acute treatment in less severe depression



Model goodness of fit and inconsistency

This section reports only findings of goodness of fit and inconsistency checks for NMA analyses that informed clinical evidence. Detailed findings of goodness of fit and inconsistency checks for all NMA analyses, including those that informed the guideline economic model are reported in the respective sections of Chapter 17.

For the SMD of depressive symptom scores, relative to the size of the intervention effect estimates, moderate to low between trial heterogeneity was observed for this outcome

Update 2017

[τ=0.22 (95% CrI 0.17 to 0.29)]. No meaningful differences were observed in posterior mean residual deviance or between study heterogeneity suggesting that there was no evidence of inconsistency.

For response in those randomised, moderate between trials heterogeneity was found relative to the size of the intervention effect estimates [τ=0.56 (95% CrI 0.45 to 0.68)]. No meaningful differences were observed in posterior mean residual deviance or between study heterogeneity suggesting that there was no evidence of inconsistency.

For remission in those randomised, moderate between trials heterogeneity was found relative to the size of the intervention effect estimates, [τ=0.49 (95% CrI 0.33 to 0.69)]. There was a substantial reduction in between study heterogeneity in the inconsistency model suggesting evidence of inconsistency. Therefore, these results should be interpreted with caution.

Detailed comparisons between the relative effects of all pairs of interventions obtained from the consistency (NMA) model and those obtained from the inconsistency (pairwise) model are provided in Appendix W.

Selective outcome reporting and publication bias

The bias adjustment models on SMD of depressive symptom scores that were developed to assess potential bias associated with small study size showed a slightly improved fit to the data compared with the unadjusted NMA, although the DIC favoured the unadjusted NMA model and there was only a small reduction in the between-study heterogeneity when adjusting for bias. The mean bias b had a negative median (as expected) but the 95% CrI included the possibility of a zero bias with moderate variability [median b=-0.22 (95% CrI -1.93 to 1.50); median standard deviation of b=0.99 (95% CrI 0.05 to 2.38)]. These findings suggest no evidence of small study bias in comparisons between active and inactive interventions in the SMD outcome.

The SMDs of classes versus pill placebo resulting from the bias adjusted model showed negligible changes in relative effects for most classes. A small reduction in effect was observed for self-help with support and long-term PDPT; however, since there was no evidence of bias these findings should be interpreted with caution. Bias adjustment had a very small impact on class rankings, which remained largely unaffected. The relative effects of all classes versus pill placebo (posterior mean SMD with 95% CrI) and posterior mean ranks of each class (with 95% CrI) obtained from the bias-adjusted model are provided, for illustrative purposes, in Table 48. Classes in the table have been ranked from smallest to largest ranking (with lower rankings suggesting better outcome). The relative effects of every class versus pill placebo obtained from the bias-adjusted model are shown in Figure 11.

For treatment discontinuation, the bias adjusted model showed a slightly improved fit to the data compared with the unadjusted NMA, although the DIC favoured the unadjusted NMA model and there was only a small reduction in the between-study heterogeneity when adjusting for bias. The mean bias b had a positive median (which is opposite to the expected direction) and the 95% CrI included the possibility of a zero bias with small variability [median b=0.18 (95% CrI -0.19 to 0.47); median standard deviation of b=0.26 (95% CrI 0.02 to 0.61)]. These findings suggest no evidence of small study bias in comparisons between active and inactive interventions in the NMA of discontinuation in those randomised.

For response in completers, the bias adjusted model showed a substantially improved fit to the data compared with the unadjusted NMA with the DIC favouring the bias adjusted NMA model. There was also a substantial reduction in the between-study heterogeneity in the bias adjusted model. The mean bias b had a positive median (as expected) and the 95% CrI excluded the possibility of a zero bias although with moderate variability [median b=1.48 (95% CrI 0.64 to 2.34); median standard deviation of b=0.68 (95% CrI 0.10 to 1.29)]. These findings provide strong evidence of small study bias in this outcome, in comparisons between

Update 2017

active and inactive interventions. For this reason, the economic analysis included a probabilistic sensitivity analysis which utilised data on response in completers derived from the bias-adjusted NMA model, to test the impact of the potential small study bias in response in completers outcome on the results of the economic analysis.

Detailed results of all bias models are provided in Appendix N, and Chapter 17, Section 17.8.

Table 48: Results of NMA bias model in people with a new episode of less severe depression. Standardised mean difference of depressive symptom scores following adjustment for small study bias: Posterior effects (SMD) of all classes versus pill placebo and ranking of classes

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Combined (IPT + AD)	63	-1.44 (-2.78 to -0.1)	2.64 (1 to 13)
Combined (Short-term PDPT + AD)	165	-1.12 (-2.47 to 0.23)	4.13 (1 to 18)
Self-help with support	514	-0.65 (-1.40 to 0.07)	6.17 (1 to 16)
Long-term PDPT	128	-0.66 (-1.54 to 0.20)	6.33 (1 to 18)
CT/CBT	2026	-0.42 (-0.79 to -0.06)	8.21 (4 to 13)
Combined (Exercise + AD/CBT)	79	-0.52 (-1.78 to 0.75)	8.49 (1 to 20)
Combined (CT/CBT + AD)	36	-0.50 (-1.83 to 0.83)	8.72 (1 to 20)
Behavioural therapies	717	-0.43 (-1.27 to 0.42)	8.83 (2 to 19)
Psychoeducational interventions	268	-0.38 (-1.19 to 0.44)	9.33 (2 to 19)
Exercise	752	-0.29 (-1.51 to 0.93)	10.73 (1 to 20)
IPT	234	-0.27 (-1.12 to 0.56)	10.76 (3 to 20)
Short-term PDPT	379	-0.22 (-0.85 to 0.42)	11.43 (4 to 19)
SSRIs	2463	-0.21 (-0.71 to 0.28)	11.44 (4 to 19)
TCAs	811	-0.20 (-0.75 to 0.35)	11.62 (4 to 19)
Self-help without support	1721	-0.15 (-0.84 to 0.55)	12.36 (4 to 20)
Counselling	406	-0.12 (-0.99 to 0.75)	12.66 (3 to 20)
Pill placebo	294	reference	14.88 (11 to 18)
Attention placebo	1577	0.04 (-0.29 to 0.38)	15.34 (10 to 19)
TAU	1675	0.12 (-0.11 to 0.36)	16.83 (13 to 19)
Waitlist	1022	0.34 (0.04 to 0.64)	19.11 (17 to 20)

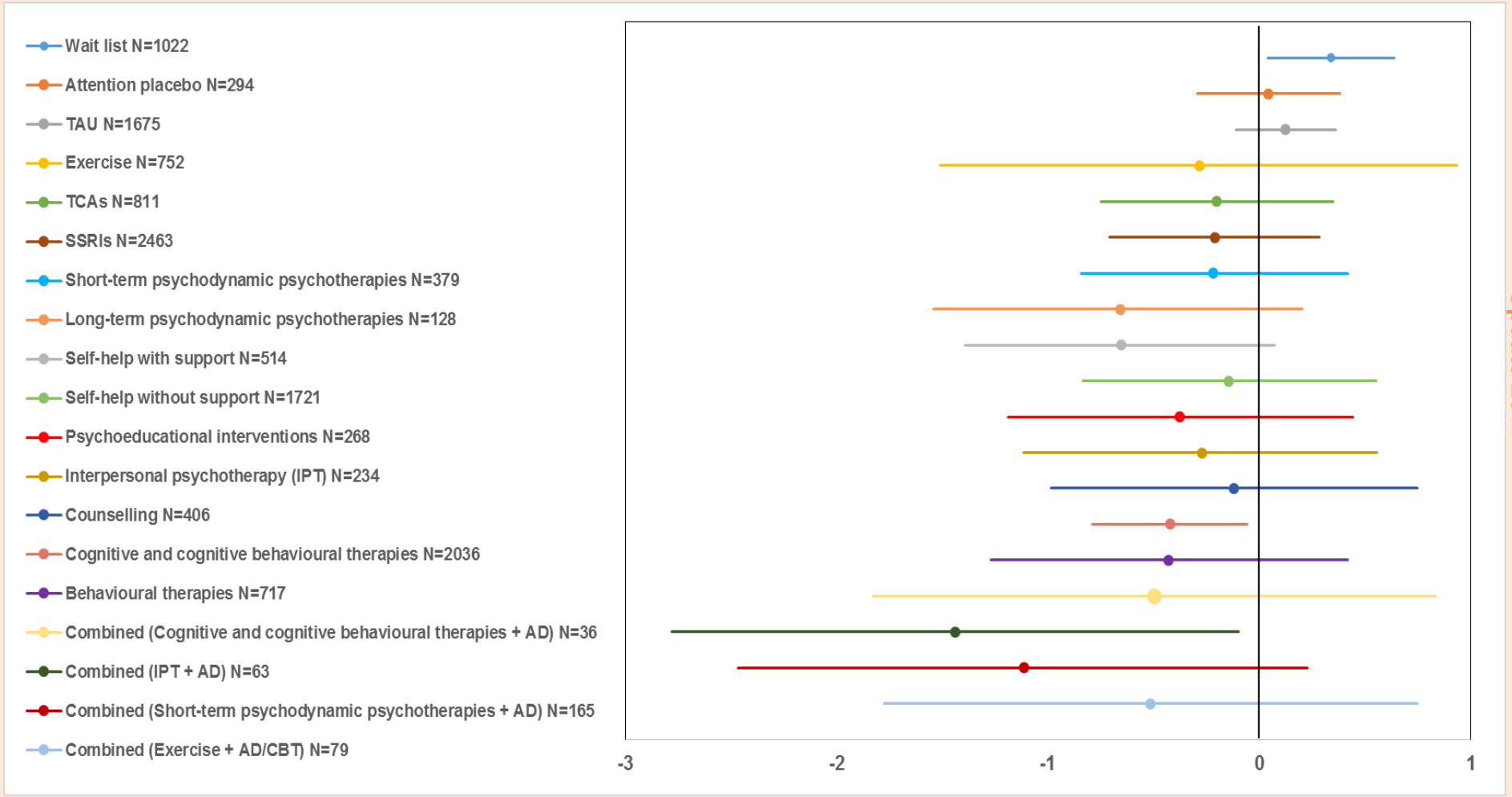
Notes:

Negative effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo)

AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants

Update 2017

Figure 11 Results of NMA bias model in people with a new episode of less severe depression. Standardised mean difference of depressive symptom score of all classes versus pill placebo (N=1564) following adjustment for small study bias [values on the left side of the vertical axis indicate a better effect compared with pill placebo]



Indirectness

In the context of the NMA, indirectness refers to potential differences across the populations, interventions and outcomes of interest, and those included in the relevant studies that informed the NMA.

A key assumption when conducting NMA is that the populations included in all RCTs considered in the NMA are similar. However, it is noted that participants in pharmacological and psychological trials may differ to the extent that some participants find different interventions more or less acceptable in light of their personal circumstances and preferences (so that they might be willing to participate in a pharmacological trial but not a psychological one and vice versa). Similarly, self-help trials may recruit participants who would not seek or accept face-to-face interventions. However, a number of trials included in the NMA have successfully recruited participants who are willing to be randomised to either pharmacological or psychological intervention and to either self-help or face-to-face treatment. The NMAs have assumed that service users are willing to accept any of the interventions included in the analyses; in practice, treatment decisions may be influenced by individual values and goals, and people's preferences for different types of interventions. These factors were taken into account when formulating recommendations.

Interventions of similar type were grouped in classes following GC advice and considered in class models. These models allowed interventions within each class to have similar, but not identical, effects around a class mean effect. Classes and interventions assessed in the NMAs were directly relevant to the classes and interventions of interest.

Outcomes reported in included studies were also the primary outcomes of interest, as agreed by the GC.

Update 2017

7.4.2 Economic evidence

7.4.2.1 Economic literature review

The systematic search of the literature identified 12 UK studies that assessed the cost effectiveness of interventions for adults with a new episode of less severe depression (Brabyn et al. 2016, Chalder et al. 2012, Kaltenthaler et al. 2006, Kendrick et al. 2005 and 2006a, Kendrick et al. 2009, Kendrick et al. 2006b and Peveler et al. 2005, Littlewood et al. 2015, McCrone et al. 2004, Phillips et al. 2014, Richards et al. 2016, Simpson et al. 2003, Spackman et al. 2014). Details on the methods used for the systematic search of the economic literature, including inclusion criteria for each review question, are described in Chapter 3. Full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix Q. Completed methodology checklists of the studies are provided in Appendix P. 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 R.

Categorisation of the studies by their population's severity level of depressive symptoms followed the same criteria used for the categorisation of the clinical studies included in the guideline systematic review. All economic studies adopted a NHS perspective, with some studies including personal social service (PSS) costs as well; in addition, some studies reported separate analyses that adopted a societal perspective. NHS and PSS cost elements included, in the vast majority of studies, intervention, primary and community care, staff time (such as GPs, nurses, psychiatrists, psychologists), medication, inpatient and outpatient care and other hospital care. All studies used national unit costs; in some studies, intervention costs were based on local prices or prices provided by the manufacturers (e.g. in the case of computerised CBT packages).

7.4.2.1.11 Psychological interventions

Problem solving

Kendrick and colleagues (2005 and 2006a) evaluated the cost effectiveness of problem-solving treatment provided by mental health nurses compared with generic community mental health nurse care and usual GP care in adults with a new episode of anxiety, depression or reaction to life difficulties, with duration of symptoms between 4 weeks to 6 months, in the UK. The economic analysis was conducted alongside a RCT (Kendrick2006, N=247; analysis based on n=184 with clinical data available; cost data available for n=159). Most of the study participants (75%) had a diagnosis of depression. The measure of outcome was the QALY, estimated based on EQ-5D ratings (UK tariff). The time horizon of the analysis was 26 weeks.

Under a NHS perspective, problem solving and generic mental health nurse care were found to be significantly more expensive than GP care. The number of QALYs gained was practically the same across all interventions, meaning that GP care was the dominant option. The study is directly applicable to the NICE decision-making context and is characterised by minor limitations.

Psychodynamic counselling

Simpson and colleagues (2003) assessed the cost effectiveness of psychodynamic counselling provided by trained, BAC accredited counsellors, who received regular supervision, in addition to usual GP treatment, versus usual GP treatment alone, in adults with depression, with or without comorbid anxiety, in the UK. The economic analysis was performed alongside of a RCT (Simpson2003, N=145; cost and outcome data at 12 months available for n=115). The outcome measure of the analysis was the change in the BDI score, with secondary outcomes including changes in scores on other scales, such as the Brief Symptom Inventory (BSI), the Inventory for Interpersonal Problems (IIP), the Social Adjustment Schedule (SAS), the Duke Social Support Scale (DSSS), plus the number of 'cases of depression' defined as BDI≥14 or any of total BSI measures ≥63, or any SAS subcategory ≥2. The duration of the analysis was 12 months.

Using a health and social services perspective, the analysis showed that psychodynamic counselling has similar costs and outcomes with usual GP treatment. Although bootstrapping was conducted to estimate uncertainty around costs and outcomes, there was no attempt to combine costs and outcomes in a single measure of cost effectiveness (ICER). The study is only partially applicable to the NICE decision-making context (as the QALY was not the measure of outcome) and is characterised by potentially serious limitations, mainly the lack of providing a summary measure of cost effectiveness that would allow a clearer conclusion on the cost effectiveness of psychodynamic counselling (and on the underlying uncertainty) to be made.

Computerised CBT (with minimal support)

McCrone and colleagues (2004) evaluated the cost effectiveness of computerised CBT (Beating the Blues package) versus treatment as usual, in adults with a diagnosis of depression, mixed depression and anxiety or anxiety disorders, alongside a RCT (Proudfoot 2004a, N=274, cost data available for n=261) that was conducted in the UK. The outcome measures used were the BDI, the number of depression-free days (DFDs) defined based on BDI scores, and the QALY that was estimated assuming that a DFD scores 1 and a day with depression scores 0.59. The time horizon of the analysis was 8 months.

Using a NHS perspective, computerised CBT was found to be more costly and more effective than treatment as usual, with ICERs of £17 per point improvement on BDI, £2 per extra DFD and £1,944 per QALY (2015 prices). The probability of computerised CBT being

Update 2017

cost-effective was 0.99 at a cost effectiveness threshold of £23,324 per QALY, which suggests that computerised CBT is likely a cost-effective intervention. However, estimation of QALYs is based on assumptions and does not follow NICE recommended methodology. The study is thus only partially applicable to the NICE decision-making context and is characterised by potentially serious limitations.

Kaltenthaler and colleagues (2006) undertook decision-analytic economic modelling to assess the cost-utility of computerised CBT versus treatment as usual in adults with depression attending primary care services in the UK. The study evaluated 3 different computerised CBT packages (Beating the Blues; Cope; Overcoming Depression). Efficacy data were taken from analysis of RCT individualised data, other published RCT data and further assumptions. Resource use data were based on manufacturer submissions, published data and other assumptions. The outcome measure was the QALY, based on EQ-5D ratings (UK tariff). The time horizon of the analysis was 18 months.

Based on a NHS perspective, computerised CBT was more costly and more effective than treatment as usual, with an ICER ranging from £2,470 to £9,791 per QALY (depending on package, uplifted to 2015 prices). The probability of computerised CBT being cost-effective ranged from 0.54 to 0.87 at a cost effectiveness threshold of £41,000 per QALY, suggesting that computerised CBT may overall be a cost-effective intervention. The study is directly applicable to the NICE decision-making context but is characterised by potentially major limitations as a number of input parameters were based on assumptions.

Computerised CBT with support

Littlewood and colleagues (2015) conducted an economic analysis alongside a RCT (Gilbody 2015, N=691; at 24 months EQ-5D data available for n=416 and NHS cost data available for n=580) to assess the cost effectiveness of 2 computerised CBT programmes with therapist support (the commercially produced package Beating the Blues and the free to use package MoodGYM) versus treatment as usual in adults with depression in the UK. The outcome measure was the QALY estimated based on EQ-5D ratings (UK tariff). The duration of the analysis was 2 years.

Update 2017

Using a NHS and PSS perspective, the commercially produced computerised CBT was more expensive than treatment as usual, and the freely available computerised CBT was less costly than treatment as usual. Treatment as usual produced a higher number of QALYs than either of the 2 computerised CBT packages. Thus, the commercially produced computerised CBT was dominated by treatment as usual. The ICER of treatment as usual versus the free-to-use computerised CBT package was £7,193 per QALY (2015 prices). The probability of treatment as usual being cost-effective across the 3 treatment options was 0.55 at the lower NICE cost effectiveness threshold of £20,000 per QALY. Using QALYs generated based on the SF-6D, the commercially produced computerised CBT programme was still dominated by treatment as usual; in contrast, the freely available computerised CBT programme became the dominant option; under this scenario, the probability of the freely available computerised CBT programme being cost effective at the lower NICE cost effectiveness threshold became 0.76. Results were robust to inclusion of depression-related costs only and to consideration of completers' data only (instead of imputed data analysis). Moreover, there was little evidence of an interaction effect between preference and treatment allocation on outcomes. These results suggest that computerised CBT with support is unlikely to be cost-effective within the NICE decision-making context (which recommends use of EQ-5D for generation of QALYs). The study is directly applicable to the UK context and is characterised by minor limitations.

Phillips and colleagues (2014) undertook an economic analysis alongside a RCT (Phillips 2014, N=637; for the clinical analysis, completion was 56% at 6 weeks and 36% at 12 weeks; for the cost analysis, completion rates were not reported) to estimate the cost effectiveness of computerised CBT with support (the freely available package of MoodGYM) versus

attention control in adults with depression in the UK. The outcome measures were the change in Work and Social Adjustment Scale (WSAS) scores and the QALY, estimated based on EQ-5D (UK tariff). The time horizon of the analysis was 12 weeks for the outcomes and 6 weeks for costs.

The time horizon of the analysis was very short and different for costs and outcomes, with very low completion rates for outcome data both at 6 and 12 weeks. Attention control was shown to be more costly and more effective than computerised CBT, with an ICER of £4,000/QALY. The study is characterised by inadequate reporting of results; no incremental analysis was conducted (although it is possible to conduct from reported data) and no uncertainty results were presented. Finally, it is unclear if the intervention cost (in terms of equipment and overheads required) has been considered in the analysis. Therefore, although the study is directly applicable to the UK context, it is characterised by very serious limitations and therefore was not further considered when formulating recommendations.

Computerised CBT with support versus computerised CBT

Brabyn and colleagues (2016) evaluated the cost effectiveness of telephone-facilitated computerised CBT compared with minimally supported computerised CBT for adults with depression in the UK. The economic analysis was conducted alongside a pragmatic multicentre RCT (Brabyn 2016, N=369; complete cost data across the trial period available for n=209). In both arms, a freely available computerised CBT program was used (MoodGYM). The outcome measure of the analysis was the QALY, estimated based on EQ-5D (UK tariff). The time horizon of the analysis was 12 months.

Under a NHS and PSS perspective, and after adjusting for baseline costs and EQ-5D score, age, anxiety level, baseline depression severity, depression duration and sex, telephone-facilitated computerised CBT was dominant over minimally supported computerised CBT (that is, it was more effective and less costly). The probability of telephone-facilitated computerised CBT being the cost-effective option was 0.55 at both the lower and the upper NICE cost effectiveness thresholds of £20,000 and £30,000 per QALY, respectively. Results were robust to inclusion of mental health-related costs only. The study is directly applicable to the NICE decision-making context and is characterised by minor limitations.

Update 2017

Behavioural activation versus cognitive behavioural therapy (CBT)

Richards and colleagues (2016) compared the cost effectiveness between behavioural activation and CBT in adults with depression in the UK, alongside a non-inferiority RCT (Richards 2016, N=440; costs available for n=327; QALYs available for n=309). The outcome measure of the analysis was the QALY, estimated based on EQ-5D ratings (UK tariff). The time horizon of the analysis was 18 months.

Under a NHS and PSS perspective, behavioural activation was dominant over CBT (that is, it was more effective and less costly). The probability of behavioural activation being the cost-effective option was 0.8 at both the lower and the upper NICE cost effectiveness thresholds of £20,000 and £30,000 per QALY, respectively. Results were robust to imputation of missing data. The study is directly applicable to the NICE decision-making context and is characterised by minor limitations.

7.4.2.1.2 Pharmacological interventions

Kendrick and colleagues (2009) evaluated the cost effectiveness of provision of SSRIs (fluoxetine, fluvoxamine, sertraline, paroxetine, citalopram or escitalopram) in addition to supportive care provided by GPs compared with GP supportive care alone in adults with mild or moderate depression in the UK. The economic analysis was conducted alongside a RCT (Kendrick 2009, N=220; 12-week completers n=196; 6-month followed-up n=160). The measures of outcome were the change in HAMD17 score and the QALY, estimated based on SF-36/SF-6D ratings (UK tariff). The time horizon of the analysis was 12 and 26 weeks.

Under a NHS and social care perspective, SSRI plus supportive care was dominant over supportive care alone at 12 weeks (i.e. it was more effective and had lower total costs). At 26 weeks, SSRI plus supportive care was still more effective but also more costly than supportive care alone, with an ICER of £106 per unit of improvement on HAMD17 or £17,429 per QALY (2015 prices). SSRI plus supportive care had a probability of being cost-effective of more than 0.50 when the cost effectiveness threshold exceeded £94 per unit reduction on HAMD17. At the NICE cost effectiveness threshold of £20,000-£30,000 /QALY, the probability of SSRI plus supportive care reached 0.65-0.75. The study is directly applicable to the NICE decision-making context and is characterised by minor limitations.

Peveler and colleagues (2005) and Kendrick and colleagues (2006b) evaluated the cost effectiveness of provision of TCAs (amitriptyline, dothiepin or imipramine), SSRIs (fluoxetine, sertraline or paroxetine) and lofepramine (a TCA that was considered in a separate arm) in adults with a new episode of mild-to-moderate depression willing to receive antidepressant treatment in primary care in the UK. The economic analysis was conducted alongside an open-label RCT with a partial preference design: following randomisation, treatment could be prescribed from a different class to the one allocated at random, if participants or their doctor preferred an alternative (Peveler 2005; N=327; entered preference group n=92; followed-up at 12 months n=171). The measures of outcome were the number of depression-free weeks (DFWs, defined as a HADS-D score <8) and the QALY based on EQ-5D ratings (UK tariff). The time horizon of the analysis was 12 months.

Under a NHS perspective, SSRIs were more costly and more effective than TCAs and lofepramine. Using the number of DFWs as the measure of outcome, TCAs were extendedly dominated (i.e. they were less effective and more expensive than a linear combination of the other 2 options). The ICER of SSRI versus lofepramine was £45 per extra DFW. Using the QALY as the measure of outcome, lofepramine was extendedly dominated. The ICER of SSRIs versus TCAs was £3,821/QALY (2015 prices). The probability of SSRIs being cost-effective was approximately 0.6 at the NICE lower cost effectiveness threshold of £20,000/QALY. The study is directly applicable to the NICE decision-making context and is characterised by minor limitations.

Update 2017

7.4.2.1.3 Physical interventions

Acupuncture versus counselling versus usual care

Spackman and colleagues (2014) evaluated the cost effectiveness of acupuncture versus counselling versus treatment as usual in adults with depression, who were in contact with primary care services for this reason in the past 5 years, in the UK. The analysis was conducted alongside an open parallel-arm RCT (MacPherson 2013, N=755; at 12 months EQ-5D data available for n=572; complete resource use data for n=150; multiple imputation used). The intervention cost of acupuncture was taken from published data, as no NHS data were available. The outcome measure of the analysis was the QALY, estimated based on EQ-5D ratings (UK tariff). The time horizon of the analysis was 12 months.

Using a NHS perspective, acupuncture was found to be the most cost-effective intervention with an ICER versus treatment as usual of £4,731/QALY (2015 prices). Counselling was extendedly dominated, with an ICER versus acupuncture of £74,449/QALY. However, the analysis indicated that when acupuncture is not an option, then counselling is cost-effective versus treatment as usual, with an ICER of £8,233/QALY. Probabilistic sensitivity analysis showed that the probability of cost effectiveness at the NICE lower cost effectiveness threshold of £20,000/QALY was 0.62 for acupuncture, 0.36 for counselling and only 0.02 for treatment as usual. Results were sensitive to small changes in intervention costs and robust to inclusion of depression-related resource use only. Using a complete case analysis acupuncture dominated counselling. The study is directly applicable to the NICE decision-making context but is characterised by potentially serious limitations, including the

particularly high proportion of missing resource use data and the sensitivity of the results to intervention costs.

Physical exercise programme

Chalder and colleagues (2012) assessed the cost effectiveness of a physical activity intervention delivered by a physical activity facilitator in addition to usual GP care versus usual GP care alone in adults with a recent first or new depressive episode in the UK. The analysis was conducted alongside a RCT, which was excluded from the clinical analysis due to high attrition rates (N=361; at 12 months EQ-5D data n=195; complete resource use data n=156; multiple imputation used in sensitivity analysis). The outcome measure of the analysis was the QALY, estimated based on EQ-5D (UK tariff). The time horizon of the analysis was 12 months.

Under a NHS and PSS perspective and using only completers' data, the physical activity intervention was found to be more costly and more effective than usual GP care, with an ICER of £22,871/QALY (2015 prices). Its probability of being cost-effective at the NICE lower (£20,000/QALY) and higher (£30,000/QALY) cost effectiveness threshold was 0.49 and 0.57, respectively. Using imputed data, the ICER of the physical activity programme versus usual GP care was £21,290/QALY, while its probability of being cost-effective at the NICE lower and higher cost effectiveness threshold rose just at 0.50 and 0.60, respectively. The study is directly applicable to the NICE decision-making context but is characterised by potentially serious limitations, mainly its notably high attrition rates.

7.4.2.21 Guideline economic modelling

Update 2017

A decision-analytic model was developed to assess the relative cost effectiveness of pharmacological, psychological, physical and combined interventions for the treatment of a new episode of less severe depression in adults. The objective of economic modelling, the methodology adopted, the results and the conclusions from this economic analysis are described in detail in Chapter 14. This section provides a summary of the methods employed and the results of the economic analysis.

Overview of economic modelling methods

A hybrid decision-analytic model consisting of a decision-tree followed by a three-state Markov model was constructed to evaluate the relative cost effectiveness of a range of pharmacological, psychological, physical and combined interventions for the treatment of a new episode of less severe depression in adults treated in primary care. The time horizon of the analysis was 12 weeks of acute treatment (decision-tree) plus 2 years of follow-up (Markov model). The interventions assessed were determined by the availability of efficacy and acceptability data obtained from the NMAs that were conducted to inform this guideline. Specific interventions were used as exemplars within each class, so that results of interventions can be extrapolated, with some caution, to other interventions of similar resource intensity within their class. The following interventions [in brackets the classes they belong to] were assessed:

- pharmacological interventions: citalopram [SSRIs]; mirtazapine [mirtazapine]
- psychological interventions: behavioural activation (BA) [behavioural therapies]; Coping with Depression course (group) [behavioural therapies]; cognitive behavioural therapy (CBT) individual (over 15 sessions) [CT/CBT]; CBT group (under 15 sessions) [CT/CBT]; interpersonal psychotherapy (IPT) [IPT]; short term psychodynamic psychotherapy (PDPT) individual [short-term PDPT]; non-directive counselling [Counselling]; computerised CBT with support [self-help with support]; computerised CBT without support [self-help without or with minimal support]; psychoeducational group programme [psychoeducational interventions]
- physical interventions: physical exercise programme [exercise]

- combined interventions: CBT individual (over 15 sessions) + citalopram [Combined CT/CBT and antidepressant]; IPT + citalopram [Combined IPT and antidepressant]; short term PDPT individual + citalopram [Combined short-term PDPT and antidepressant]; physical exercise programme + sertraline [Combined exercise and CBT or antidepressant]
- clinical management, reflecting GP visits, corresponding to pill placebo RCT arms.

The decision-tree component model structure considered the events of discontinuation for any reason and specifically due to intolerable side effects; treatment completion and response reaching remission; treatment completion and response not reaching remission; treatment completion and inadequate or no response. The Markov component model structure considered the states of remission, depressive episode (due to non-remission or relapse), and death. The specification of the Markov component of the model was based on the relapse prevention model developed for this guideline, details of which are provided in Chapter 13.

Efficacy data were derived from the guideline systematic review and NMAs. Baseline parameters (baseline risk of discontinuation, discontinuation due to side effects, response in treatment completers and remission) were estimated based on a review of naturalistic studies. The measure of outcome of the economic analysis was the number of QALYs gained. Utility data were derived from a systematic review of the literature, and were generated using EQ-5D measurements and the UK population tariff. The perspective of the analysis was that of health and personal social care services. Resource use was based on published literature, national statistics and, where evidence was lacking, the GC expert opinion. National UK unit costs were used. The cost year was 2016. Model input parameters were synthesised in a probabilistic analysis. This approach allowed more comprehensive consideration of the uncertainty characterising the input parameters and captured the non-linearity characterising the economic model structure. A number of one-way deterministic sensitivity analyses were also carried out. In addition, a probabilistic sensitivity analysis that used data on response in completers derived from NMAs adjusted for bias resulting from small study size (as described in Section 7.3.6) was undertaken.

Results have been expressed in the form of Incremental Cost Effectiveness Ratios (ICERs) following the principles of incremental analysis. Net Monetary Benefits (NMBs) have also been estimated. Incremental mean costs and effects (QALYs) of each intervention versus clinical management (pill placebo) have been presented in the form of cost effectiveness planes. Results of probabilistic analysis have been summarised in the form of cost effectiveness acceptability curves (CEACs), which express the probability of each intervention being cost effective at various cost effectiveness thresholds). Moreover, cost effectiveness acceptability frontiers (CEAFs) have also been plotted; these show the treatment option with the highest mean NMB over different cost effectiveness thresholds, and the probability that the option with the highest NMB is the most cost-effective among those assessed.

Overview of economic modelling results and conclusions

In people with less severe depression, pharmacological treatment, group psychological interventions and other low-intensity psychological and physical interventions were the most cost-effective options. These were followed by high intensity psychological interventions alone or in combination with pharmacological treatment, a number of which appeared to be less cost-effective than clinical management. The ranking of interventions, from the most to least cost-effective, was as follows: mirtazapine, CBT group, physical exercise programme, citalopram (representing SSRIs), cCBT with support (representing self-help with support), physical exercise programme combined with sertraline, psychoeducational group programme, Coping with Depression group course (representing behavioural therapies delivered in groups), cCBT without or with minimal support (representing self-help without or with minimal support), CBT individual, BA (representing individual behavioural therapies), IPT combined with citalopram, clinical management by GPs (reflecting pill placebo trial

Update 2017

arms), IPT, short term PDPT individual, short term PDPT individual combined with citalopram (or another antidepressant), counselling, CBT individual combined with citalopram (or another antidepressant). The probability of mirtazapine being the most cost-effective option was 0.45 at the NICE lower cost effectiveness threshold of £20,000/QALY.

Results of the economic analysis were overall robust to different scenarios explored through sensitivity analysis. The relative cost effectiveness of high intensity psychological interventions, alone or combined with antidepressants, improves when these are delivered by less specialised therapists, such as Band 5 psychological well-being practitioners -PWPs- or Band 6 therapists (instead of Band 7 clinical psychologists) and deteriorates when higher utility values are assumed at baseline, as the scope for HRQoL improvement following successful treatment is more limited, and when a 50% lower cost of relapse is assumed at baseline. The cost effectiveness of counselling improves if it is delivered in 8 instead of 16 sessions.

Conclusions from the guideline economic analysis refer mainly to people with depression who are treated in primary care for a new depressive episode; however, they may be relevant to people in secondary care as well, given that clinical evidence was derived from a mixture of primary and secondary care settings (however, it needs to be noted that costs utilised in the guideline economic model were mostly relevant to primary care).

Results need to be interpreted with caution due to the limited evidence base characterising some of the interventions assessed in the models and methodological limitations characterising some of the NMAs that were used to populate the economic analyses. In particular, data were limited for more than one outcomes for mirtazapine and IPT combined with citalopram. In addition, the NMA on remission in completers was characterised by inconsistency between direct and indirect evidence.

Update 2017

7.4.3 Clinical evidence statements

- Evidence from 63 randomised participants suggests a large and statistically significant benefit of a combined IPT and antidepressant intervention relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the highest ranked intervention for clinical efficacy as measured by SMD of depressive symptom scores (mean rank 2.71, 95% CrI 1 to 13).
- Evidence from 165 randomised participants suggests a large but not statistically significant benefit of a combined short-term psychodynamic psychotherapy and antidepressant intervention relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the second highest ranked intervention for clinical efficacy as measured by SMD (mean rank 4.23, 95% CrI 1 to 17).
- Evidence from 514 randomised participants suggests a large and statistically significant benefit of self-help with support relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the third highest ranked intervention for clinical efficacy as measured by SMD (mean rank 5.48, 95% CrI 1 to 14).
- Evidence from 128 randomised participants suggests a moderate to large, but not statistically significant, benefit of long-term psychodynamic psychotherapy relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the fourth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 5.99, 95% CrI 1 to 17).
- Evidence from 79 randomised participants suggests a moderate but not statistically significant benefit of a physical exercise programme combined with CBT or an antidepressant relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the fifth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 8.11, 95% CrI 1 to 20).
- Evidence from 2026 randomised participants suggests a small to moderate and statistically significant benefit of a cognitive or cognitive behavioural intervention relative to

pill placebo on depression symptomatology for adults with less severe depression, and this was the sixth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 8.16, 95% CrI 4 to 13).

- Evidence from 36 randomised participants suggests a moderate but not statistically significant benefit of a cognitive or cognitive behavioural intervention combined with an antidepressant relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the seventh highest ranked intervention for clinical efficacy as measured by SMD (mean rank 8.68, 95% CrI 4 to 13).
- Evidence from 717 randomised participants suggests a small to moderate, but not statistically significant, benefit of a behavioural therapy relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the eighth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 9.11, 95% CrI 2 to 19).
- Evidence from 268 randomised participants suggests a small and not statistically significant benefit of a psychoeducational intervention relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the ninth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 10.05, 95% CrI 2 to 20).
- Evidence from 752 randomised participants suggests a small and not statistically significant benefit of a physical exercise programme relative to pill placebo on depression symptomatology for adults with less severe depression, and this was the tenth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 10.37, 95% CrI 1 to 20).
- Evidence from 234 randomised participants suggests a small and not statistically significant benefit of IPT relative to pill placebo on depression symptomatology for adults with less severe depression, and this intervention was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 10.81, 95% CrI 3 to 20).
- Evidence from 2463 randomised participants suggests a small and not statistically significant benefit of an SSRI relative to pill placebo on depression symptomatology for adults with less severe depression, and this intervention was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 11.38, 95% CrI 4 to 19).
- Evidence from 379 randomised participants suggests a small and not statistically significant benefit of short-term psychodynamic psychotherapy relative to pill placebo on depression symptomatology for adults with less severe depression, and this intervention was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 11.51, 95% CrI 4 to 19).
- Evidence from 811 randomised participants suggests a small and not statistically significant benefit of a TCA relative to pill placebo on depression symptomatology for adults with less severe depression, and this intervention was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 11.93, 95% CrI 5 to 19).
- Evidence from 1721 randomised participants suggests a small and not statistically significant benefit of self-help without support relative to pill placebo on depression symptomatology for adults with less severe depression, and this intervention was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 12.02, 95% CrI 4 to 20).
- Evidence from 406 randomised participants suggests no benefit of counselling relative to pill placebo on depression symptomatology for adults with less severe depression, and this intervention was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 12.71, 95% CrI 3 to 20).

Update 2017

- Evidence from 294 randomised participants suggests no difference between attention-placebo relative to pill placebo on depression symptomatology for adults with less severe depression, and both control interventions were ranked alongside each other for clinical efficacy as measured by SMD (mean rank 14.90, 95% CrI 11 to 18 for pill placebo relative to 15.09, 95% CrI 10 to 19, for attention placebo).
- Evidence from 1675 randomised participants suggests no difference between treatment as usual relative to pill placebo on depression symptomatology for adults with less severe depression, and this intervention was ranked second from bottom for clinical efficacy as measured by SMD (mean rank 17.48, 95% CrI 14 to 19).
- Evidence from 1035 randomised participants suggests a lower effect of wait list compared with pill placebo on depression symptomatology for adults with less severe depression; the difference in effect was small and not statistically significant. Waitlist was ranked bottom for clinical efficacy as measured by SMD (mean rank 19.31, 95% CrI 17 to 20).

7.4.4 Economic evidence statements

7.4.4.1 Psychological interventions

- Evidence from 1 single UK study conducted alongside a RCT (N = 247) suggests that problem solving is unlikely to be cost-effective compared with treatment as usual in adults with a new episode of less severe depression. The evidence is directly applicable to the UK context and is characterised by minor limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N = 145) is inconclusive as to whether psychodynamic counselling is cost-effective in adults with a new episode of less severe depression. The evidence is partially applicable to the NICE decision-making context and is characterised by potentially serious limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N = 274) and 1 study based on economic modelling suggests that computerised CBT (with minimal support) may be potentially cost-effective compared with treatment as usual in adults with a new episode of less severe depression. The evidence comes from a directly applicable (model-based) study and a partially applicable (RCT-based) study and is characterised by potentially serious limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N = 691) indicates that computerised CBT with support is unlikely to be cost-effective compared with treatment as usual in adults with a new episode of less severe depression. The evidence is directly applicable to the UK context and is characterised by minor limitations. Evidence from another single study conducted alongside a RCT (N=637) indicates that computerised CBT with support is unlikely to be cost-effective compared with attention control. The evidence is directly applicable to the UK context but is characterised by very serious limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N = 369) indicates that computerised CBT with support may be cost-effective compared with computerised CBT with minimal support in adults with a new episode of less severe depression. The evidence is directly applicable to the UK context and is characterised by minor limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N = 440) indicates that behavioural activation is likely to be cost-effective compared with CBT in adults with less severe depression. The evidence is directly applicable to the UK context and is characterised by minor limitations.

7.4.4.2 Pharmacological interventions

- Evidence from 1 single UK study conducted alongside a RCT (N = 220) indicates that provision of SSRIs in addition to GP supportive care is likely to be cost-effective compared with GP supportive care alone in adults with a new episode of less severe depression.

Update 2017

The evidence is directly applicable to the UK context and is characterised by minor limitations.

- Evidence from 1 single UK study conducted alongside an open label RCT with a partial preference design (N = 327; entering preference group n=92) indicates that provision of SSRIs is likely to be more cost-effective than TCAs or lofepramine in adults with a new episode of less severe depression. The evidence is directly applicable to the UK context and is characterised by minor limitations.

7.4.4.38 Physical interventions

- Evidence from 1 single UK study conducted alongside a RCT (N = 755) indicates that acupuncture is likely to be cost-effective compared with counselling and treatment as usual in adults with a new episode of less severe depression. The evidence is directly applicable to the UK context but is characterised by potentially serious limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N = 361) suggests that a physical exercise programme is potentially cost-effective compared with treatment as usual in adults with a new episode of less severe depression. The evidence is directly applicable to the UK context but is characterised by potentially serious limitations.

7.4.4.47 Pharmacological, psychological, physical and combined interventions

- Evidence from the guideline economic modelling suggests that pharmacological treatment, group psychological therapies (such as group CBT) and other low-intensity psychological and physical interventions are the most cost-effective options for the treatment of new episodes of less severe depression in adults. High-intensity psychological interventions appear to be less cost-effective. CBT individual, BA (representing individual behavioural therapies) and IPT combined with citalopram (or another antidepressant) appear to be more cost-effective than clinical management (comprising GP visits) whereas IPT alone, short-term PDPT individual alone or combined with citalopram (or another antidepressant), counselling, and CBT individual combined with citalopram (or another antidepressant) appear to be less cost-effective than clinical management. This evidence refers mainly to people treated in primary care for a new depressive episode; however, it may be relevant to people treated in secondary care as well, given that clinical evidence was derived from a mixture of primary and secondary care settings. The economic analysis is directly applicable to the NICE decision-making context and is characterised by minor limitations, although the evidence base for some interventions is rather limited, and respective results should therefore be interpreted with caution.

Update 2017

7.4.5 From evidence to recommendations

7.4.5.1 Relative values of different outcomes

The GC used the results of economic modelling (cost effectiveness) as the main criterion for making recommendations and the NMA results on the SMD of depressive symptom scores outcome (ranking of interventions and relative effects versus pill placebo) as a secondary criterion. Economic modelling was informed by a range of outcomes of the NMAs (discontinuation for any reason, discontinuation due to side effects, response in completers, remission in completers) but not by the SMD outcome. The GC used pill placebo as a benchmark in both the clinical and economic analyses and expressed the view that for an intervention to be recommended, it should show higher cost effectiveness and a better clinical effect compared with pill placebo.

7.4.5.21 Trade-off between clinical benefits and harms

In developing the recommendations in this guideline the GC were mindful of a number of important factors which underpin the effective delivery of care for people with depression and the need to ensure that medication is properly monitored and reviewed, paying attention to the reduction of potential harms. The GC agreed that not addressing these factors could lead to poorer engagement with the service, higher attrition, sub-optimal delivery of treatments and consequent poorer outcomes. The GC therefore developed a number of recommendations, based on their informal consensus, which required all interventions to be provided in the context of effective assessment, care planning, liaison and outcome monitoring; to use appropriate manuals and competence frameworks supported by effective supervision and audit to support the effective implementation of interventions.

In relation to medication, the GC were concerned that the recommendations developed for this guideline stressed the importance of fully informing service users about the benefits and potential harms of medication (including discontinuation symptoms and how they might be managed), the importance of continuing with the agreed dose and of gradually reducing the dose when stopping medication. The GC also thought it important to be clear about the management of suicide risk particularly in younger people and the toxicity associated with certain medication (in particular with tricyclic antidepressants). The GC recognised the increased side effect burden with certain drugs in particular lithium and antipsychotic medication and therefore decided to make a recommendation on the physical health care monitoring of people taking these drugs as they were concerned that the SPCs for these drugs are not always followed. The GC's purpose in developing these recommendations was to reduce potential harm that may occur and also to increase uptake of and reduce attrition rates for what are helpful interventions

The GC were predominantly guided by the results of the health economic analysis for those interventions covered by the NMA when drafting the recommendations for people with less severe depression. These recommendations were supported by a review of the relative effectiveness of the interventions against pill placebo

The GC reviewed the rankings of all interventions and noted the ranking of the 6 most effective classes of interventions based on the SMD of depressive symptom scores outcome were combined interpersonal therapy + antidepressants, combined short-term psychodynamic psychotherapy + antidepressants, self-help with support, long-term psychodynamic psychotherapy, combined exercise + antidepressant or cognitive behavioural therapy, and cognitive and cognitive behavioural therapy. For the 3 clinical outcomes assessed (SMD of depressive symptom scores, response in those randomised and remission in those randomised) the rankings of the classes that ranked in the top six places are summarised below:

- combined interpersonal therapy with antidepressants and combined short-term psychodynamic psychotherapy with antidepressants were in the top six rankings for all 3 outcomes;
- cognitive and cognitive behavioural therapies, combined exercise with antidepressants or with cognitive behavioural therapy, and long-term psychodynamic psychotherapy were in the top six rankings for 2 of the outcomes;
- behavioural therapies, counselling, interpersonal psychotherapy, combined cognitive and cognitive behavioural therapies with antidepressant, self-help with support and mirtazapine were in the top six rankings for 1 outcome.

The GC noted that the inclusion of classes in the top six rankings was affected by data availability. Mirtazapine was in the top six rankings only for the outcome of response in those randomised; however, this was the only outcome for which mirtazapine data were available. Similarly, long-term psychodynamic psychotherapy was in the top 6 classes for the outcomes of SMD and response in those randomised, but was not included in the remission in those

Update 2017

randomised analysis due to lack of relevant data. All other classes had available data that were included across the three analyses.

The GC also took into account that there would need to be some flexibility in the treatment options for people with less severe depression, to enable both service user choice and availability of alternative treatment options dependant on past experience of treatment or tolerability problems.

For all severities of depression, the GC agreed that the likely benefits of the recommendations made would be improvements in depression symptoms, remission and response. The potential harms identified were attrition, not taking up of other treatments, issues with acceptability (particularly for drugs which have more side effects) and the possibility of people deteriorating (as data in clinical trials of all treatments estimated this could happen in 7-10% of people). In developing the recommendations the GC also took into account the harm-to-benefit ratio of antidepressants and how the balance of harm and benefit would vary with different severities of depression

7.4.5.35 Trade-off between net health benefits and resource use

Existing economic evaluations assessed a limited range of pharmacological, psychological and physical interventions in, mostly, pairwise comparisons, so it was difficult for the GC to draw any robust conclusions on the relative cost effectiveness of the full range of interventions that are available for the treatment of adults with a new episode of less severe depression.

The guideline economic analysis assessed the cost effectiveness of a wide range of pharmacological, psychological, physical and combined interventions, as well as clinical management (GP visits, reflected in pill placebo trial arms) as initial treatments for people with a new episode of less severe depression. The interventions included in the economic analysis were dictated by availability of data and were used as exemplars within their class, as for practical reasons it was impossible to model all interventions considered in the guideline NMA. Therefore the GC noted that results of interventions could be extrapolated, with some caution, to other interventions of similar resource intensity within the same class.

The GC based the guideline recommendations primarily on the findings of the guideline economic analysis. The ranking of interventions for adults with a new episode of less severe depression, from the most to the least cost-effective was: mirtazapine, CBT group, physical exercise programme, citalopram, cCBT with support, physical exercise programme combined with sertraline, psychoeducational group programme, Coping with Depression course (group), cCBT without or with minimal support, CBT individual, behavioural activation, IPT combined with citalopram, clinical management, IPT, short term psychodynamic psychotherapy individual, short term psychodynamic psychotherapy individual combined with citalopram, counselling, and CBT individual combined with citalopram. The GC considered the probabilities of cost effectiveness obtained using a step-wise approach, according to which the most cost-effective intervention is omitted at each step and the probability of the next most cost-effective intervention is re-calculated and the uncertainties around cost effectiveness.

The GC took into account the strengths and the limitations of the economic analysis, the robustness of the results under different scenarios explored through sensitivity analysis (including use of data from the NMA bias models), and noted that a number of interventions in the economic analysis were informed by limited data or borrowed efficacy from a different intervention, in particular mirtazapine and IPT combined with citalopram.

Based on the above considerations, the GC decided to recommend group CBT as a first option for the treatment of new episodes of less severe depression in adults, as it was the second most cost-effective intervention and belonged to a class (cognitive and cognitive behavioural therapies) with a robust evidence base and a high ranking on the SMD outcome.

Update 2017

The GC decided to recommend individual self-help with support as an alternative for people who do not want group CBT, because it was the next most cost-effective psychological intervention and ranked very highly on the SMD outcome.

The GC recommended a physical activity programme for people with less severe depression who do not want, group CBT or self-help with support because it ranked in third place in cost effectiveness.

The GC made a 'consider' recommendation for mirtazapine and SSRIs (represented by citalopram in the economic analysis) based on previous treatment history or the person's preference due to the limited evidence available for mirtazapine (limited data informing the economic analysis and no data available on the SMD outcome), the relatively low ranking of SSRIs on the SMD outcome, and the harm-to-benefit ratio of antidepressants in a population with less severe depression in combination with the availability of other cost-effective treatment options.

The GC acknowledged the lower cost effectiveness of high intensity individual psychological interventions compared with low intensity psychological interventions and drugs, but expressed the opinion that some of these interventions may be suitable options for people with a history of poor response to psychological or pharmacological interventions in a previous episode of depression or a history of good response to specific high intensity psychological interventions or a potential risk of developing more severe depression. The GC also noted that the economic analysis assumed that all individual psychological interventions are delivered by a Band 7 clinical psychologist and that their relative cost effectiveness improved if these were effectively delivered by therapists paid at a lower Band.

After reviewing the cost effectiveness results and the clinical results on the SMD outcome, the GC decided to recommend individual CBT or behavioural activation for these populations, as both interventions appeared to be more cost-effective than clinical management and belonged to classes with the highest ranking (and a robust evidence base) on the SMD outcome among classes of psychological interventions.

The GC noted that, although long-term psychodynamic psychotherapy ranked in a higher place than CBT and behavioural therapies, this was not included in the economic analysis due to lack of suitable data, but, nevertheless, it was very unlikely to be cost-effective, given its high resource use intensity.

The GC considered the marginally lower cost effectiveness of IPT compared with clinical management (ICER of IPT versus clinical management £22,612/QALY) and decided to make a 'consider' recommendation for IPT in people with less severe depression for whom other recommended interventions (group CBT, physical activity programme, facilitated self-help, pharmacological interventions, individual CBT or BA) had not worked well in a previous episode of depression or in those who did not want the other recommended interventions and who would like help for interpersonal difficulties that focus on role transition, disputes or grief. The GC expressed the view that the effectiveness and cost effectiveness of IPT was likely to be higher in this sub-population compared with the 'general' population with less severe depression that was the focus of the guideline economic analysis.

The GC considered the lower cost effectiveness of counselling compared with clinical management (ICER of counselling versus clinical management £25,913/QALY) and decided to make a 'consider' recommendation for counselling in people with less severe depression for whom other recommended interventions (group CBT, physical activity programme, facilitated self-help, pharmacological interventions, individual CBT or BA) had not worked well in a previous episode of depression or in those who did not want the other recommended interventions and who would like help for significant psychosocial, relationship or employment problems. The GC expressed the view that the effectiveness and cost effectiveness of counselling was likely to be higher in this sub-population compared with the 'general' population with less severe depression that was the focus of the guideline economic

Update 2017

analysis. The GC also noted that according to the guideline economic analysis the cost effectiveness of counselling improved when this was effectively delivered by therapists paid at Band 6 or when this was delivered in 8 sessions, and agreed that these scenarios tested in sensitivity analysis may comprise variations of clinical practice in some settings.

The GC considered the lower cost effectiveness of short-term psychodynamic psychotherapy compared with clinical management (ICER of short-term psychodynamic psychotherapy versus clinical management £25,441/QALY) and decided to make a 'consider' recommendation for short-term psychodynamic psychotherapy in people with less severe depression for whom other recommended interventions (group CBT, physical activity programme, facilitated self-help, pharmacological interventions, individual CBT or BA) had not worked well in a previous episode of depression or in those who did not want the other recommended interventions and who would like help for emotional and developmental difficulties in relationships. The GC expressed the view that the effectiveness and cost effectiveness of short-term psychodynamic psychotherapy was likely to be higher in this sub-population compared with the 'general' population with less severe depression that was the focus of the guideline economic analysis.

The GC were concerned that psychological interventions are not always implemented consistently – for example audits have suggested that reduced numbers of sessions are used in practice compared with what is recommended. They therefore agreed it was important to specify the structure of the psychological interventions being recommended to ensure consistency. The recommended structure of all psychological interventions (number and duration of sessions, number of therapists and participants for group interventions) was based on the resource use utilised in the economic analysis, which, in turn, was informed by RCT resource use, modified by the GC expert advice to represent routine clinical practice in the UK, so that recommended structure of psychological interventions represents cost-effective use of available healthcare resources as implemented in routine clinical practice.

Update 2017

7.4.5.27 Quality of evidence

The GC took into account that evidence for some treatments on the SMD outcome was limited (people randomised in combined CT/CBT + antidepressant N=36; combined IPT + antidepressant N=63; combined short-term psychodynamic psychotherapy + antidepressant N=165; long-term psychodynamic psychotherapy N=128; combined exercise and antidepressant/CBT N=79) and non-existent for mirtazapine. Among psychological treatments, CT/CBT had the most robust evidence base (N=2,026; mean effect versus pill placebo -0.46, 95% CrI -0.85 to -0.07); among pharmacological treatments, SSRIs had the most robust evidence base (N=2,463; mean effect versus pill placebo -0.23, 95% CrI -0.89 to 0.45). It was noted that there was no evidence of inconsistency for the SMD outcome. However, there was evidence of inconsistency for the remission in those randomised, as well as for the remission in completers outcome; the latter informed the economic analysis.

The bias adjustment models on SMD suggested no evidence of small study bias in comparisons between active and inactive interventions. For outcomes used in the economic analysis, there was strong evidence of small study bias in response in completers, in comparisons between active and inactive interventions. However, the GC noted that a sensitivity analysis of the economic model showed economic results to be robust to bias adjustment.

Overall, the GC considered that the quality of the evidence, both clinical and economic, was robust enough to allow recommendations to be based on the available evidence, although in forming recommendations they took into account that the evidence on mirtazapine was quite thin.

7.4.5.51 Other considerations

The GC wanted to compare the findings of the guideline NMAs with those of published reviews and meta-analyses of psychological interventions for people with depression. They noted the different methodology adopted for the guideline NMAs compared with published reviews, which could justify potential differences in results: the guideline NMAs included well-defined populations, without physical comorbidities, who were treated for a new episode of depression; 2 NMAs were conducted separately for people with less severe and people with more severe depression. An important difference between the guideline NMAs and published reviews (including published NMAs) was the inclusion of drug and self-help trials in the analysis. Interventions included in the guideline NMAs were defined and classified differently from other reviews. The guideline NMAs utilised class models, where individual treatment effects are drawn towards a class mean but individual intervention estimates are retained and are more precise. The evidence base used for each NMA analysis was broader than in other reviews, with a combination of continuous (including change from baseline, use of baseline and endpoint mean scores) and dichotomous data being used to inform the SMD and response analyses; a hierarchy of depressive symptom scales was used for this purpose, following GC expert advice.

The GC inspected comparisons between active classes included in the NMA and noted that the results of the NMAs for people with less severe depression are broadly consistent with those of published reviews.

The GC noted, based on the evidence that where there was no or limited facilitation of computerized CBT there was an increased rate of attrition from the interventions. Therefore the GC decided to emphasize the importance of facilitation in delivering a range of self-help interventions, including computerized interventions.

The GC discussed the issue of patient choice, with the lay members offering the opinion that many people are happy solely with a choice of either evidence based psychological or pharmacological therapy, with choices between different therapies of the same modality being of less concern. They thought that there would be a subset of patients who would have researched therapies carefully and would have a strong preference, but that this would not apply to the majority of people. Other issues such as choice of the gender of the therapist, the setting in which interventions were provided and good information on the content of, potential harms or side effects and likely outcomes of an intervention were also considered important.

In developing these recommendations, the GC considered the relative training, experience and salaries of staff providing a range of psychological interventions including counselling, behavioural activation and cognitive behavioural therapies. The GC were aware of the different levels of experience and salary of therapists in some of the trials which form the evidence base. However, the GC took the view that as the majority of high intensity therapists were paid either at AfC Grade 6 or 7 that it was appropriate to use these salaries for the base case economic analysis.

Update 2017

7.4.6 Recommendations

General principles of care

All interventions

33. Support people with depression to decide on their preferences for interventions by giving them:

- information on what the interventions are, and the expected outcomes

- choice on the intervention type, how it will be delivered (face to face or digitally), and where it will be delivered
- the option, if possible, to choose the gender of the practitioner
- information on what the next steps will be if the initial intervention is not helpful.

34. Provide interventions for people with depression in a framework. This should include:

- an assessment of need
- the development of a treatment plan
- taking into account any physical health problems
- regular liaison between healthcare professionals in specialist and non-specialist settings
- routine outcome monitoring and follow-up. [new 2017]

35. Use psychological and psychosocial treatment manuals[d] to guide the form and length of the intervention [2017].

36. Consider using competence frameworks developed from treatment manual(s) for psychological and psychosocial interventions to support effective training delivery and supervision of interventions. [2017]

37. For all interventions for people with depression:

- use sessional outcome measures
- review how well the treatment is working with the person
- monitor and evaluate treatment adherence. [2017]

38. Healthcare professionals delivering interventions for people with depression should:

- receive regular high-quality supervision
- have their competence monitored and evaluated, for example, by using video and audio tapes, and external audit. [2017]

Pharmacological interventions

39. When offering a person antidepressant medication:

- explain the reasons for offering it
- discuss the risks and benefits
- discuss any concerns they have about taking the medication
- ensure they have information to take away that is appropriate for their needs. [2017]

40. When prescribing antidepressant medication, give people information about:

- how long it takes (typically 2-4 weeks) to begin to start to feel better
- how important it is to follow the instructions on when to take antidepressant medication
- how treatment might need to carry on even after remission

Update 2017

[d] Treatment manuals that have evidence for their efficacy from clinical trials are preferred.

- how they may be affected when they first start taking antidepressant medication, and what these effects might be
- how they may be affected if they have to take antidepressant medication for a long time and what these effects might be, especially in people over 65
- how taking antidepressant medication might affect their sense of resilience (how strong they feel and how well they can get over problems) and being able to cope
- how taking antidepressant medication might affect any other medicines they are taking
- how they may be affected when they stop taking antidepressant medication, and how these effects can be minimised
- the fact that they cannot get addicted to antidepressant medication. [2017]

41. **Advise people taking antidepressant medication that although it is not addictive, if they stop taking it, miss doses or don't take a full dose, they may have discontinuation symptoms such as:**
    - more mood changes
    - restlessness
    - problems sleeping
    - unsteadiness
    - sweating
    - abdominal symptoms
    - altered sensations.

    Explain that these discontinuation symptoms are usually mild and go away after a week but can sometimes be severe, particularly if the antidepressant medication is stopped suddenly. [2017]

42. **When stopping an antidepressant medication, slowly reduce the dose based on how long the person has been taking it. For example:**
    - over several days if the person has been taking it for 2-8 weeks
    - over several weeks if the person has been taking it for 2-12 months
    - over several months if the person has been taking it for 12 months or more. [new 2017]

43. **If a person has discontinuation symptoms when they stop taking antidepressant medication or lower their dose, reassure them that they are not having a relapse of their depression. Explain that:**
    - these symptoms are common
    - relapse does not usually happen as soon as you stop taking an antidepressant or lower the dose
    - even if they start taking an antidepressant medication again or increase their dose, the symptoms won't go away immediately. [new 2017]

44. **If a person has mild discontinuation symptoms when they stop taking antidepressant medication:**
    - monitor their symptoms
    - keep reassuring them that such symptoms are common. [new 2017]

Update 2017

45. **If a person has severe discontinuation symptoms, consider restarting the original antidepressant medication at the dose that was previously effective, or another antidepressant from the same class with a longer half-life. Reduce the dose gradually while monitoring symptoms. [new 2017]**

46. **When prescribing antidepressant medication for people with depression who are under 30 years or are thought to be at increased risk of suicide:**
    - see them 1 week after starting the medication
    - review them frequently until the risk of suicide is reduced. [2017]

47. **Take into account toxicity in overdose when prescribing an antidepressant medication for people at significant risk of suicide. Be aware that:**
    - tricyclic antidepressants (TCAs), except lofepramine, are associated with the greatest risk in overdose
    - compared with other equally effective antidepressant medication recommended for routine use in primary care, venlafaxine is associated with a greater risk of death from overdose. [2017]

48. **When prescribing antidepressant medication for older people (65 years and over):**
    - consider prescribing them at a lower dose
    - take into account the person's general physical health and possible interactions with any other medicines they may be taking
    - carefully monitor the person for side effects. [2017]

49. **For people with depression taking lithium, monitor:**
    - renal and thyroid function and calcium levels before treatment and every 6 months during treatment, or more often if there is evidence of renal impairment
    - serum lithium levels 1 week after starting treatment and at each dose change until stable, and every 3 months after that. [2017]

50. **Consider ECG monitoring in people taking lithium who have a high risk of cardiovascular disease. [2017]**

51. **For people with depression who are taking an antipsychotic[e], monitor and review:**
    - weight, initially and then weekly for the first 6 weeks, then at 12 weeks, at 1 year and then annually (plotted on a chart)
    - lipid and glucose levels at 12 weeks, at 1 year and then annually
    - adverse effects, for example, extrapyramidal side effects and prolactin-related side effects with risperidone. [2017]

52. **Do not routinely provide medication management on its own as an intervention for people with depression. [new 2017]**

Update 2017

[e] At the time of consultation (July 2017), antipsychotics 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 Prescribing guidance: prescribing unlicensed medicines for further information. See individual SPCs for full list of monitoring requirements.

First line treatment for less severe depression

53. **Offer group-based cognitive behavioural therapy (CBT) specific to depression as the initial treatment for people with less severe depression. [new 2017]**

54. **Deliver group-based CBT that is:**
    - based on a cognitive and behavioural model
    - delivered by 2 competent practitioners
    - consists of up to 9 sessions of 90 minutes each, for up to 12 participants
    - takes place over 12–16 weeks, including follow-up. [new 2017]

55. **Offer individual self-help with support for people with less severe depression who do not want group CBT. [new 2017]**

56. **Follow the principles of CBT when providing self-help with support. It should:**
    - provide age-appropriate, written, audio or digital (computer or online) material
    - have support from a trained practitioner who facilitates the self-help intervention, encourages completion and reviews progress and outcome
    - consist of up to 6 sessions (face-to-face or by telephone or online), each up to 30 minutes
    - take place over 9–12 weeks, including follow-up. [2017]

57. **Consider a physical activity programme specifically designed for people with depression who do not want group CBT or self-help with support. [new 2017]**

58. **Ensure physical activity programmes for people with less severe depression:**
    - are delivered in groups by a competent practitioner
    - consist of 45 minutes of aerobic exercise of moderate intensity and duration twice a week for 5 weeks, then once a week for a further 7 weeks
    - usually have 8 people per group. [new 2017]

59. **Consider a selective serotonin reuptake inhibitor (SSRI) or mirtazapine for people who with less severe depression who choose not to have psychological interventions or based on previous treatment history for confirmed depression had a positive response to SSRIs or mirtazapine or had a poor response to psychological interventions. [new 2017]**

60. **Offer individual CBT or behavioural activation (BA) if a person with less severe depression:**
    - has a history of poor response when they tried group CBT, a physical activity programme, facilitated self-help or antidepressant medication before **or**
    - has responded well to CBT or BA before **or**
    - is at risk of developing more severe depression, for example they have a history of severe depression or the current assessment suggests a more severe depression is developing. [new 2017]

Update 2017

61. **Consider interpersonal therapy (IPT) if a person with less severe depression would like help for interpersonal difficulties that focus on role transitions or disputes or grief and:**
    - has had group CBT, exercise or facilitated self-help, antidepressant medication, individual CBT or BA for a previous episode of depression, but this did not work well for them, **or**
    - does not want group CBT, exercise or facilitated self-help, antidepressant medication, individual CBT or BA. [new 2017]

62. **Provide individual CBT, BA or IPT to treat less severe depression over 16 sessions, each lasting 50-60 minutes, over 3-4 months. [new 2017]**

63. **When giving individual CBT, BA or IPT, also consider providing:**
    - 2 sessions per week for the first 2-3 weeks of treatment for people with less severe depression
    - 3-4 follow-up and maintenance sessions over 3-6 months after finishing the course for all people who have had individual CBT, BA or IPT. [new 2017]

64. **Consider counselling if a person with less severe depression would like help for significant psychosocial, relationship or employment problems and:**
    - has had group CBT, exercise or facilitated self-help, antidepressant medication, individual CBT or BA for a previous episode of depression, but this did not work well for them, **or**
    - does not want group CBT, exercise or facilitated self-help, antidepressant medication, individual CBT or BA. [new 2017]

65. **Ensure counselling for people with less severe depression:**
    - is based on a model developed specifically for depression
    - consists of up to 16 individual sessions each lasting up to an hour
    - takes place over 12 to 16 weeks, including follow-up. [new 2017]

66. **Consider short-term psychodynamic psychotherapy (STPT) if a person with less severe depression would like help for emotional and developmental difficulties in relationships and:**
    - has had group CBT, exercise or facilitated self-help, antidepressant medication or individual CBT for a previous episode of depression, but this did not work well for them, **or**
    - does not want group CBT, exercise or facilitated self-help, antidepressant medication or individual CBT. [new 2017]

67. **Ensure STPT for people with less severe depression:**
    - is based on a model developed specifically for depression
    - consists of up to 16 individual sessions each lasting up to an hour
    - takes place over 12 to 16 weeks, including follow-up. [new 2017]

Update 2017

7.5 Review question

- For adults with a new episode of more severe depression, what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination for the treatment of depression?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 49. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 49: Clinical review protocol summary for the review of acute treatment for more severe depression

Component	Description
Review question	For adults with a new episode of more severe depression, what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination for the treatment of depression? (RQ2.2)
Population	• Adults receiving first line treatment for a new episode of depression, as defined by a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on scales (and including those with subthreshold depressive symptoms). If some, but not all, of a study's participants are eligible for the review, for instance, mixed anxiety and depression diagnoses, and we are unable to obtain the appropriate disaggregated data, then we will include a study if at least 80% of its participants are eligible for this review Baseline mean scores are used to classify study population severity according to less severe (RQ 2.1) or more severe (RQ 2.2) using the thresholds outlined in Table 38.If baseline mean scores are not available, severity will be classified according to the inclusion criteria of the study or the description given by the study authors (but only in cases where this is unambiguous, i.e. 'severe' or 'subthreshold' or 'mild').
Intervention(s)	The following interventions will be included in the NMA: **Psychological interventions:** • Behavioural therapies (including behavioural activation, behavioural therapy [Lewinsohn 1976], coping with depression course [individual and group] and social rhythm therapy [SRT]) • Cognitive and cognitive behavioural therapies (including CBT individual or group [defined as under or over 15 sessions], problem solving, rational emotive behaviour therapy [REBT] and third-wave cognitive therapies individual or group) • Counselling (including directive counselling, emotion-focused therapy [EFT], non-directive counselling and relational client-centred therapy) • Interpersonal psychotherapy • Psychodynamic psychotherapies (including individual or group-based short-term psychodynamic psychotherapy, long-term psychodynamic psychotherapy and psychodynamic counselling) • Psychoeducational interventions (including psychoeducational group programmes, intensive clinical management and lifestyle factors discussion) • Self-help with or without support (including cognitive bibliotherapy with or without support, computerised CBT [CCBT] with or without support, computerised problem solving therapy with or without support, computerised psychodynamic therapy with or without support, online positive psychological intervention and self-examination therapy) **Pharmacological interventions:** • SSRIs (citalopram, escitalopram, sertraline, fluoxetine)

Update 2017

Component	Description
	• TCAs (amitriptyline, lofepramine) • Other antidepressant drugs (mirtazapine) Note that in order to maximise connectivity in the network specific drugs that are excluded and 'any antidepressant' or 'any SSRI' or 'any TCA' nodes will be added where they have been compared against a psychological intervention and/or combined with a psychological intervention but they will not be considered as part of the decision problem. **Physical interventions:** • Exercise (including yoga) The following interventions may be compared in pairwise comparisons (however will not be included in the NMA): • Acupuncture • Behavioural couples therapy • Light therapy (for depression but not for SAD) • Nortriptyline (for older adults) • Omega-3 fatty acids • Psychosocial interventions (including befriending, mentoring, peer support and community navigators)
Comparison	• Any other intervention • Treatment as usual • Waitlist • Placebo • Imipramine
Critical outcomes	Critical outcomes **Efficacy:** • Depression symptomology (mean endpoint score or change in depression score from baseline) • Remission (usually defined as a cut off on a depression scale) • Response (e.g. reduction of at least 50% from the baseline score on HAMD/MADRS) **Acceptability/tolerability:** • Discontinuation due to side effects (for pharmacological trials) • Discontinuation due to any reason (including side effects) The following depression scales will be included in the following hierarchy: x. MADRS xi. HAMD xii. QIDS xiii. PHQ xiv. CGI xv. CES-D xvi. BDI xvii. HADS-D (depression subscale) xviii. HADS (full scale) Only one continuous scale will be used per study

Update 2017

Component	Description
	• For studies reporting response and/or remission, the scale used in the study to define cut-offs for response and/or remission will be used. • If more than one definition is used, a hierarchy of scales will be adopted (hierarchy listed above). • For studies not reporting dichotomous data, a hierarchy of scales will be adopted for continuous outcomes.
Study design	• Systematic reviews of RCTs • RCTs • Cluster RCTs

7.5.11 Clinical evidence

7.5.1.12 Study characteristics

1372 studies were considered for inclusion in this review. Of these, 145 RCTs (k=145, n=23,176) were included in this network meta-analysis.

Of the 145 RCTs included within this network and reporting either a HAM-D or MADRS score at baseline, the mean depression severity scores were HAM-D=27.3 (n=49) and MADRS=30.6 (n=26) respectively. Eight were UK based RCTs.

For a full list of included and excluded studies, study characteristics of included studies and risk of bias appendices please see Appendix J3.1 and J3.2.

Data were not available for every outcome of interest for the majority of included RCTs. For the outcomes considered in the clinical analysis, the following information was available:

- SMD of depressive symptom scores: 15 trials reported CFB data; 40 trials reported mean baseline and endpoint symptom scores and another 13 reported dichotomous response data. In total, 68 RCTs provided data on 11,300 trial participants that were used to inform the SMD outcome.
- Response in those randomised: 63 studies reported dichotomous response data, another 6 reported CFB data and in 27 studies baseline and endpoint symptom scores were available. In total, 96 RCTs with data on 15,563 participants informed this outcome.
- Remission in those randomised: 23 studies provided dichotomous remission data on 5,690 participants.

Relevant information on the number of studies and study participants that provided data on the outcomes that were used to inform the economic analysis are provided in Chapter 17, in respective outcome sections. The studies and data that were used in the NMAs for every outcome of interest are provided in Appendix T.

7.5.1.25 Results of the network meta-analysis

This section reports only NMA results that informed clinical evidence. Detailed NMA findings on all outcomes, including those that informed the economic analysis, are reported in the respective sections of Chapter 17.

Standardised mean difference (SMD) of depressive symptom scores

The network diagram of all studies included in this analysis by class is provided in Figure 12. The network diagram of the studies included in this analysis by intervention is provided in Chapter 17, Section 17.3.2.7.The relative effects of all classes versus pill placebo and versus TAU (posterior mean SMD with 95% CrI) are provided in Table 50, together with posterior

Update 2017

mean ranks of each class (with 95% CrI). Classes in the table have been ranked from smallest to largest ranking (with lower rankings suggesting better outcome). The relative effects of every class versus pill placebo and of every intervention versus pill placebo are shown in Figure 13 and Figure 14, respectively. Detailed results are provided in Chapter 17, Sections 17.3.2.7, 17.9 and 17.10.

Figure 12 Network diagram of all studies included in the analysis of standardised mean difference (SMD) of depressive symptom scores in people with a new episode of more severe depression by class

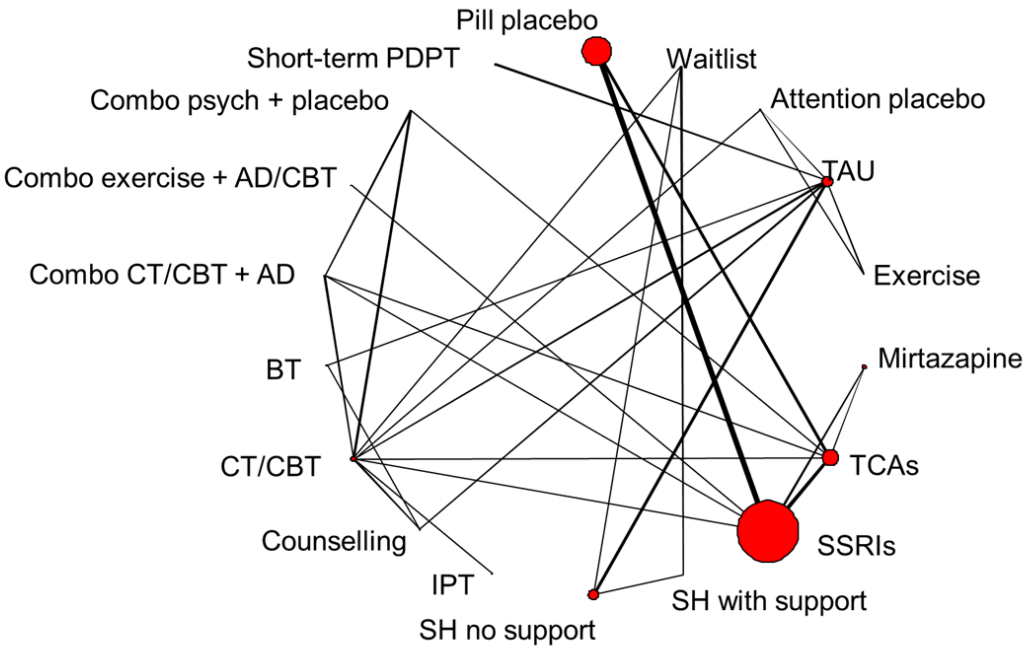


Table 50 Results of NMA in people with a new episode of more severe depression. Standardised mean difference of depressive symptom scores: Posterior effects (SMD) of all classes versus pill placebo and TAU and ranking of classes

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Effect vs TAU (mean, 95% CrI)	Mean rank (95% CrI)
Combined (Exercise + AD/CBT)	41	-1.08 (-2.55 to 0.39)	-2.30 (-4.03 to -0.59)	2.89 (1 to 10)
CT/CBT	391	-0.86 (-1.99 to 0.26)	-2.08 (-2.95 to -1.21)	3.04 (1 to 8)
TCAs	1260	-0.54 (-1.18 to 0.10)	-1.76 (-2.78 to -0.74)	4.33 (1 to 9)
Combined (CT/CBT + AD)	58	-0.52 (-1.67 to 0.62)	-1.74 (-2.91 to -0.58)	4.76 (1 to 11)
SSRIs	4696	-0.29 (-0.81 to 0.24)	-1.51 (-2.54 to -0.49)	5.90 (2 to 11)
Mirtazapine	326	-0.24 (-0.73 to 0.24)	-1.46 (-2.49 to -0.46)	6.25 (2 to 11)
IPT	95	-0.05 (-1.90 to 1.80)	-1.27 (-3.02 to 0.49)	7.41 (1 to 16)
Pill placebo	2229	reference	-1.22 (-2.15 to -0.30)	8.17 (5 to 12)
Behavioural therapies	126	0.21 (-1.70 to 2.11)	-1.01 (-2.71 to 0.67)	8.62 (1 to 17)
Self-help with support	54	0.59 (-1.51 to 2.67)	-0.63 (-2.57 to 1.33)	10.43 (1 to 17)
Counselling	120	0.69 (-0.92 to 2.34)	-0.53 (-1.95 to 0.93)	11.09 (3 to 17)
Short-term PDPT	115	0.88 (-0.93 to 2.68)	-0.34 (-1.89 to 1.20)	11.95 (3 to 17)

Update 2017

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Effect vs TAU (mean, 95% CrI)	Mean rank (95% CrI)
Self-help without support	757	0.96 (-0.49 to 2.41)	-0.26 (-1.39 to 0.87)	12.50 (6 to 17)
Exercise	50	1.04 (-0.54 to 2.60)	-0.18 (-1.48 to 1.11)	12.82 (5 to 17)
Attention placebo	80	1.16 (0.11 to 2.23)	-0.06 (-0.75 to 0.63)	13.74 (10 to 17)
TAU	825	1.22 (0.30 to 2.15)	reference	14.16 (11 to 17)
Waitlist	60	1.41 (0.24 to 2.58)	0.19 (-0.69 to 1.05)	14.93 (11 to 17)

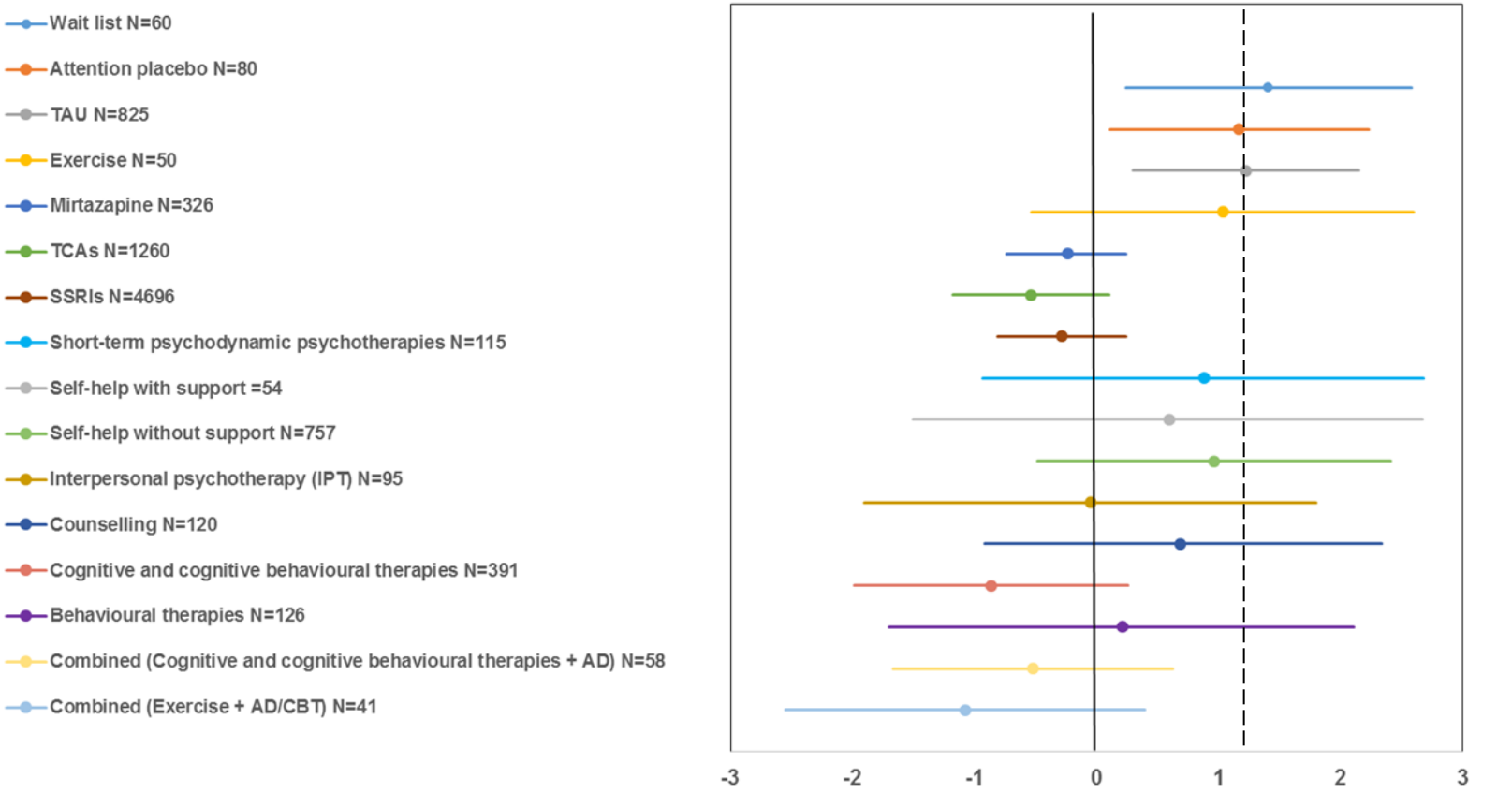
Notes:
Negative effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo or TAU)
AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants

Update 2017

1

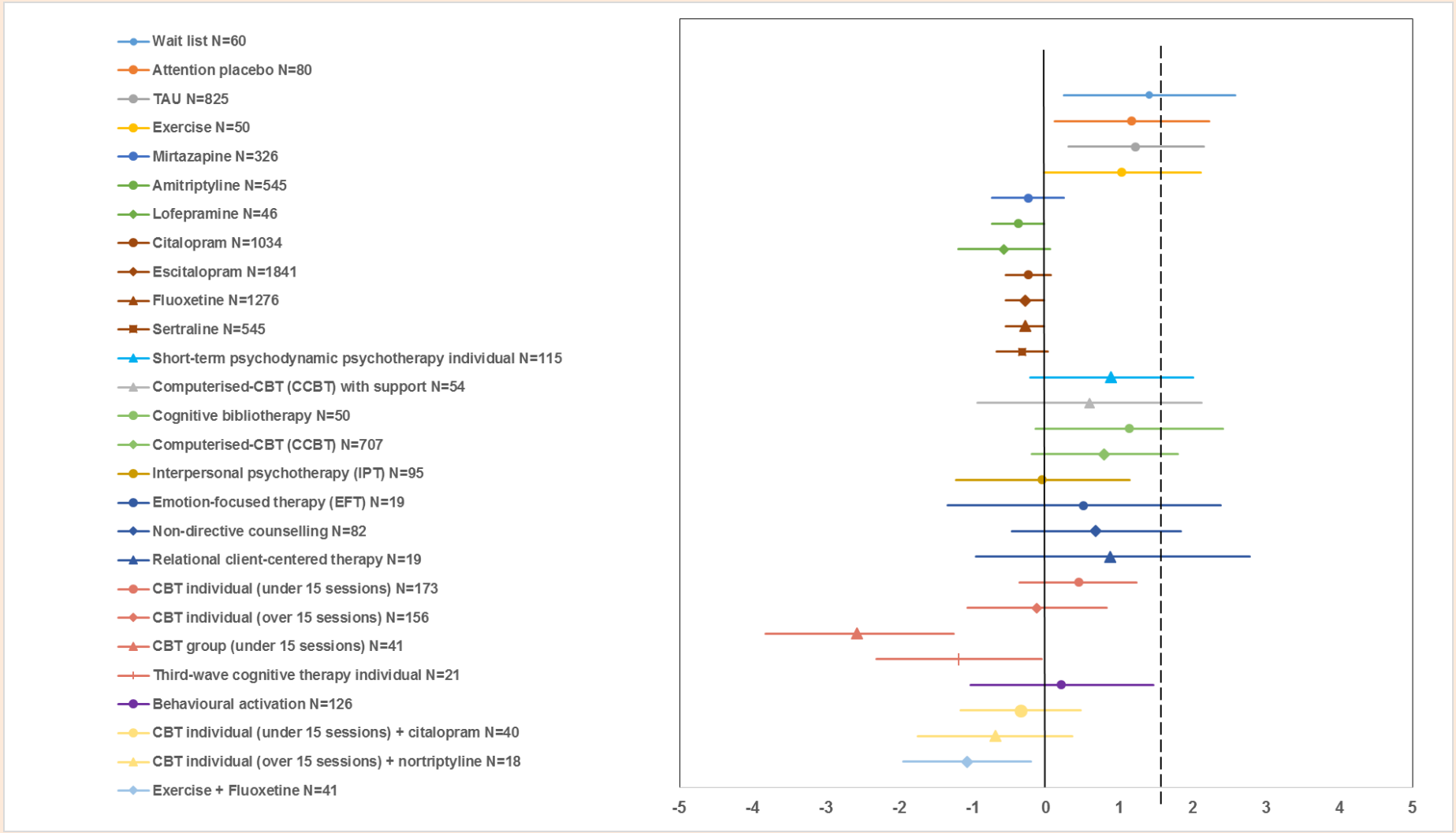
2

Figure 13 Results of NMA in people with a new episode of more severe depression. Standardised mean difference (SMD) of depressive symptom scores of all classes versus pill placebo (N=2229) [values on the left side of the vertical axis indicate a better effect compared with pill placebo; dotted line indicates TAU effect]



Update 2017

Figure 14 Results of NMA in people with a new episode of more severe depression. Standardised mean difference (SMD) of depressive symptom scores of all interventions versus pill placebo (N=2229) [values on the left side of the vertical axis indicate a better effect compared with pill placebo; dotted line indicates TAU effect]



Update 2017

Response in those randomised

The network diagram of all studies included in this analysis by class is provided in Figure 15. The network diagram of studies included in this analysis by intervention is provided in Chapter 17, section 17.3.2.6.The relative effects of all classes versus pill placebo (posterior mean LORs with 95% CrI) are provided in Table 51, together with posterior mean ranks of each class (with 95% CrI). Classes in the table have been ranked from smallest to largest ranking (with lower rankings suggesting better outcome). The relative effects of every class versus pill placebo are shown in Figure 16. Detailed results are provided in Chapter 17, section 17.3.2.6, 17.3.9 and 17.3.10.

Figure 15 Network diagram of all studies included in the analysis of response in those randomised in people with a new episode of more severe depression by class

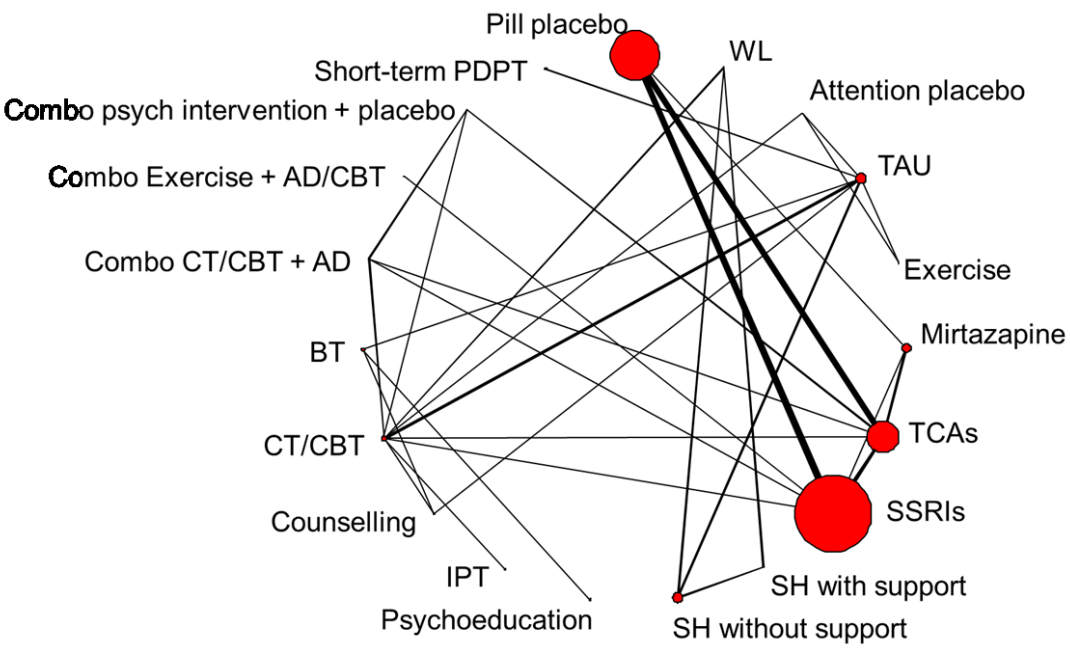


Table 51 Results of NMA in people with a new episode of more severe depression. Response in those randomised: Posterior effects (Log-Odds Ratios of response) of all classes versus pill placebo and ranking of classes

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Combined (Exercise + AD/CBT)	41	3.39 (1.35 to 5.41)	1.09 (1 to 2)
TCAs	2419	1.12 (0.37 to 1.87)	3.01 (2 to 6)
Combined (CT/CBT + AD)	58	0.97 (-0.62 to 2.59)	3.73 (1 to 8)
Mirtazapine	645	0.75 (0.21 to 1.31)	4.19 (2 to 7)
SSRIs	5874	0.56 (-0.04 to 1.15)	4.93 (2 to 8)
CT/CBT	391	-0.06 (-1.68 to 1.60)	6.57 (3 to 10)
Pill placebo	3725	reference	7.03 (5 to 10)
IPT	95	-0.79 (-3.22 to 1.64)	8.36 (2 to 14)
Behavioural therapies	236	-1.60 (-4.17 to 1.03)	10.31 (4 to 17)
Self-help with support	54	-2.34 (-4.96 to 0.28)	12.27 (7 to 18)

Update 2017

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Counselling	120	-2.50 (-4.72 to -0.28)	12.76 (8 to 18)
Short-term PDPT	120	-2.54 (-5.01 to -0.03)	12.86 (8 to 18)
Exercise	50	-2.79 (-5.07 to -0.47)	13.59 (9 to 18)
Self-help without support	757	-3.11 (-5.15 to -1.03)	14.66 (10 to 18)
Waitlist	60	-3.12 (-5.03 to -1.21)	14.78 (11 to 18)
Attention placebo	80	-3.22 (-5.12 to -1.28)	15.09 (10 to 18)
TAU	830	-3.48 (-5.07 to -1.85)	16.29 (14 to 18)

Notes:

Positive effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo)

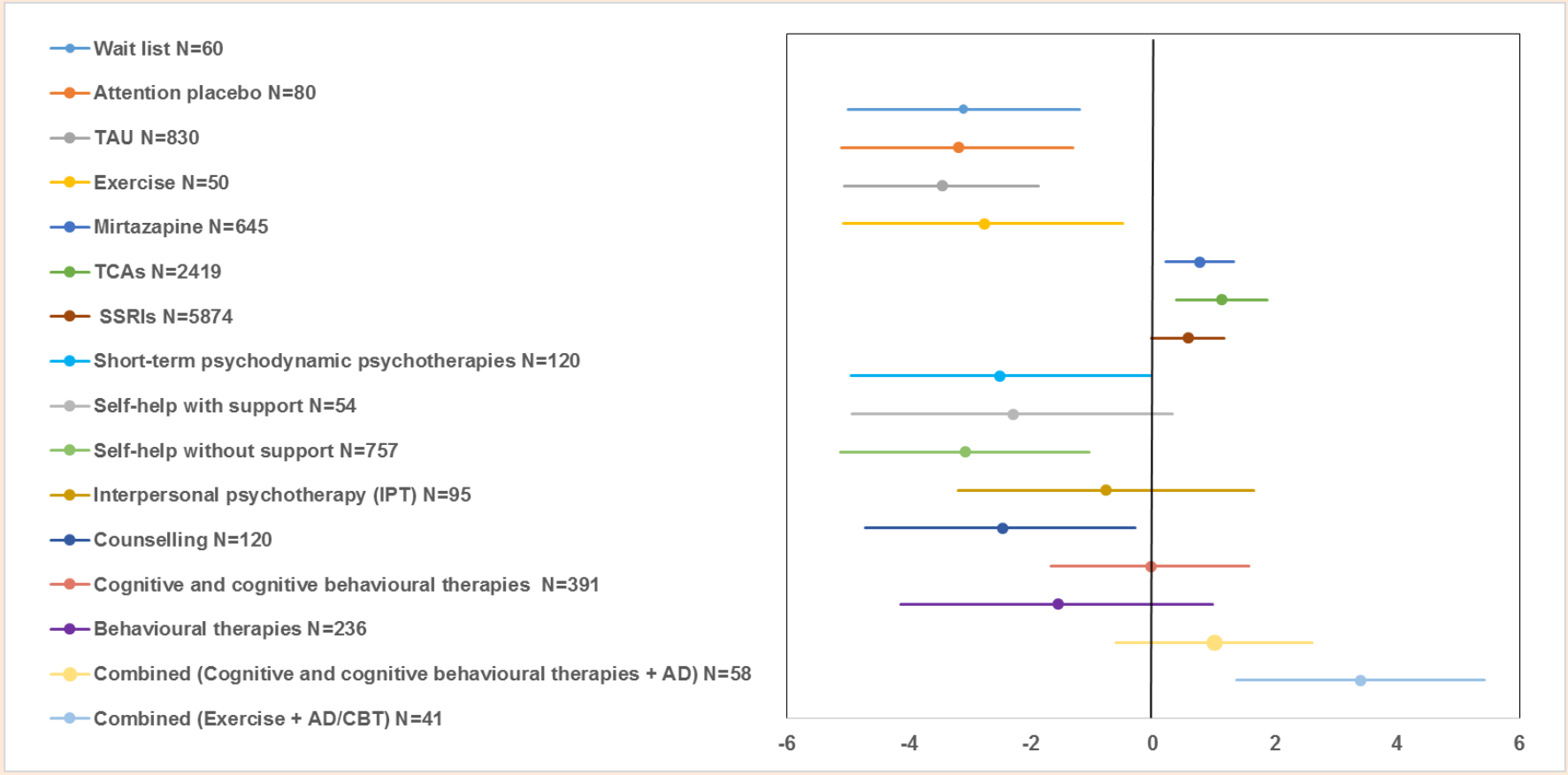
AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants

1

2

Update 2017

Figure 16 Results of NMA in people with a new episode of more severe depression. Log-Odds Ratios of response in those randomised of all classes versus pill placebo (N=3725) [values on the right side of the vertical axis indicate a better effect compared with pill placebo]



Update 2017

Remission in those randomised

The network diagram of all studies included in this analysis by class is provided in Figure 17. The network diagram of studies included in this analysis by intervention is provided in Chapter 17, Section 17.3.2.4.The relative effects of all classes versus pill placebo (posterior mean LORs with 95% CrI) are provided in Table 52, together with posterior mean ranks of each class (with 95% CrI). Classes in the table have been ranked from smallest to largest ranking (with lower rankings suggesting better outcome). The relative effects of every class versus pill placebo are shown in Figure 18. Detailed results are provided in Chapter 17, Sections 17.3.2.4, 17.9 and 17.10.

Figure 17 Network diagram of all studies included in the analysis of remission in those randomised in people with a new episode of more severe depression by class

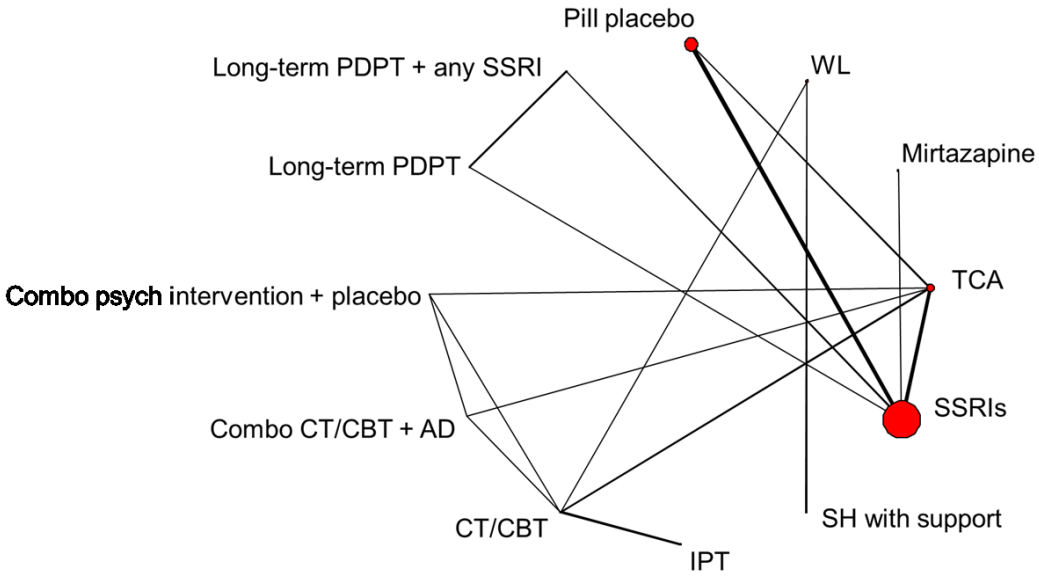


Table 52 Results of NMA in people with a new episode of more severe depression. Remission in those randomised: Posterior effects (Log-Odds Ratios) of all classes versus pill placebo and ranking of classes

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Long-term PDPT individual	90	2.46 (0.42 to 4.50)	1.86 (1 to 6)
Long-term PDPT individual + AD	91	2.05 (0.01 to 4.10)	2.45 (1 to 7)
Combined (CT/CBT + AD)	43	1.24 (-0.59 to 3.08)	3.66 (1 to 8)
CT/CBT	171	0.50 (-1.51 to 2.49)	5.63 (2 to 9)
IPT	75	0.51 (-1.20 to 3.04)	5.69 (1 to 11)
TCAs	620	0.28 (-0.83 to 1.41)	6.33 (3 to 10)
SSRIs	3097	0.18 (-0.60 to 0.96)	6.62 (3 to 10)
Mirtazapine	66	0.01 (-1.73 to 1.75)	7.08 (3 to 11)
Pill placebo	1076	Reference	7.42 (4 to 10)
Self-help with support	149	-1.63 (-6.01 to 2.40)	8.93 (2 to 11)

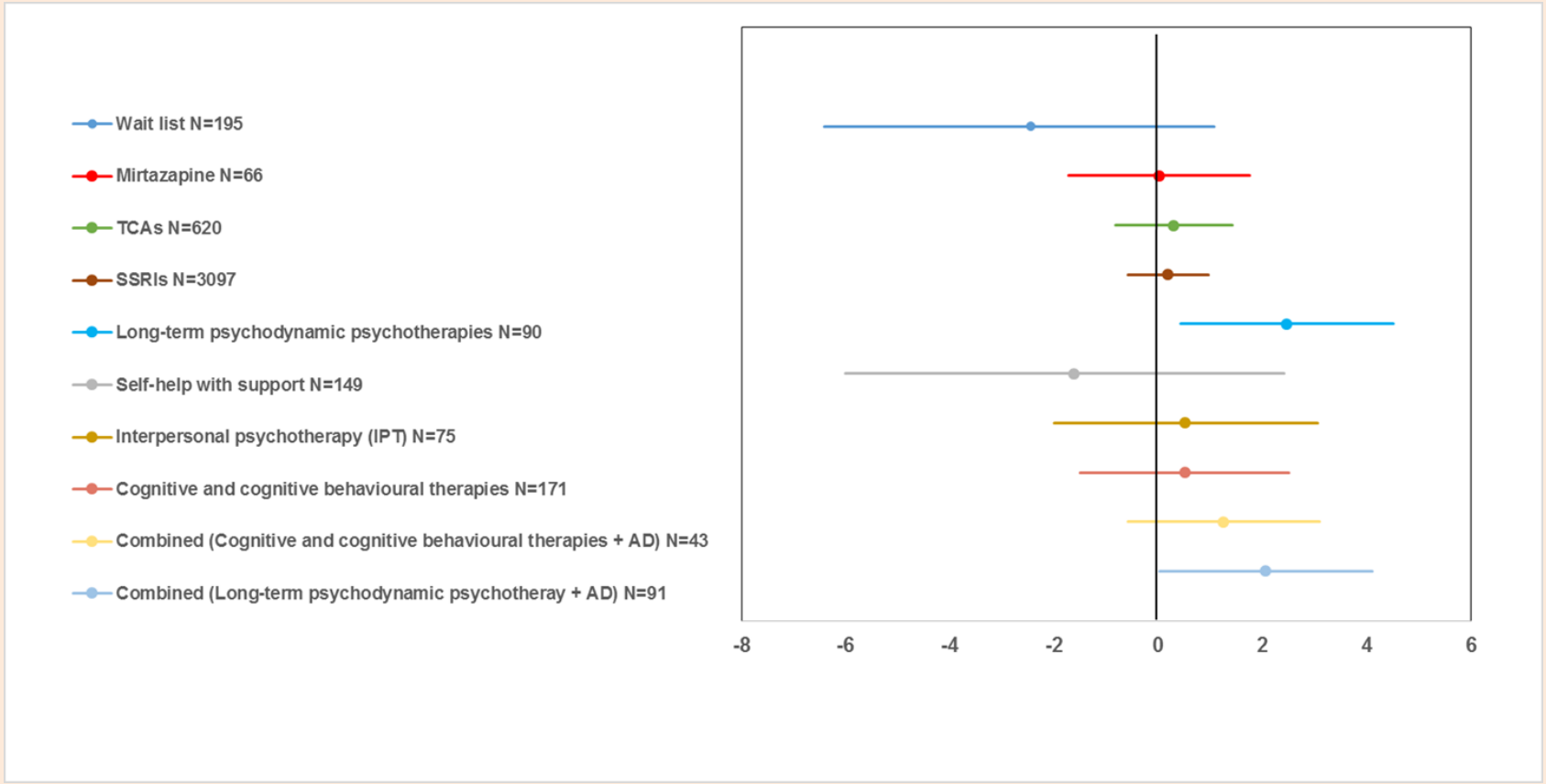
Update 2017

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Mean rank (95% CrI)
Waitlist	195	-2.44 (-6.41 to 1.07)	10.35 (6 to 11)

Notes:
Positive effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo)
AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TCAs: tricyclic antidepressants

1
2
3
4

Figure 18 Results of NMA in people with a new episode of more severe depression. Log-Odds Ratios of remission in those randomised of all classes versus pill placebo (N=1076) [values on the right side of the vertical axis indicate a better effect compared with pill placebo]



Update 2017

Comparison of the results of the NMAs that informed clinical evidence: SMD of depressive symptom scores, response in those randomised and remission in those randomised

A comparison of the results of the NMAs across the 3 outcomes of SMD of depressive symptom scores, response in those randomised and remission in those randomised can be made by inspection of Table 53. It can be seen that ranking of class effects and rankings versus pill placebo were quite consistent, in particular between the SMD and response in those randomised analyses:

- Pharmacological classes of interventions (TCAs, SSRIs and mirtazapine) showed higher effects and rankings on the SMD and response in those randomised outcomes compared with the remission in those randomised outcome, where they showed a small or no benefit compared with pill placebo.
- Self-help without or with minimal support and self-help with support showed a lower effect compared with pill placebo across all analyses (remission data were not available for self-help without or with minimal support).
- Regarding high-intensity psychological interventions, CT/CBT showed broadly consistent benefits across all analyses and had rather high rankings (2-6). IPT showed no benefit relative to pill placebo on the SMD outcome; a lower effect than pill placebo on the response in those randomised outcome; and a benefit relative to pill placebo on the remission in those randomised outcome. Behavioural therapies, counselling and short-term psychodynamic psychotherapy showed a lower effect than pill placebo in the SMD and in the response in those randomised analyses; no remission data were available for these three classes and therefore they were not included in the remission in those randomised analysis. Long-term psychodynamic psychotherapy showed a large benefit in remission in those randomised compared with pill placebo and was ranked first; however it was not included in the other two analyses due to lack of relevant data.
- Exercise showed a lower effect than pill placebo in the SMD and in the response in those randomised analyses; no remission data were available for exercise and therefore it was not included in the remission in those randomised analysis.
- Combined interventions demonstrated the highest effects and rankings. Combined CT/CBT with antidepressants was ranked in places 3-4 across the 3 analyses. Combined exercise with CBT/antidepressants ranked first in both SMD and response in those randomised analyses; it was not included in the remission in those randomised analysis due to lack of relevant data. Combined long-term psychodynamic psychotherapy with antidepressants showed a large benefit in remission in those randomised compared with pill placebo and was ranked second; however it was not included in the other two analyses due to lack of relevant data.

It needs to be noted that the 3 analyses were informed by different datasets. Nevertheless, the SMD and response in those randomised analyses may have potentially shared some study data, as in studies not reporting continuous data, dichotomous response data, if available, were used in the estimation of SMD and, conversely, in studies not reporting dichotomous response data, continuous symptom scale data, if available, were used in the estimation of response in those randomised. In contrast, the remission in those randomised analysis utilised different data from the other two analyses, which, in part, explains the inclusion of different interventions and the discrepancies observed in the results of some classes between this and the other two analyses.

Update 2017

1

Table 53 Comparison of NMA results across the outcomes considered in clinical analyses for people with a new episode of more severe depression: posterior effects of all classes versus pill placebo

Effect of every class versus pill placebo (mean, 95% CrI); classes listed according to their mean ranking (lowest to largest) for each outcome					
SMD of depressive symptom scores		**Response in those randomised (LORs)**		**Remission in those randomised (LORs)**	
Combined (Exercise + AD/CBT)	-1.08 (-2.55 to 0.39)	Combined (Exercise + AD/CBT)	3.39 (1.35 to 5.41)	Long-term PDPT individual	2.46 (0.42 to 4.50)
CT/CBT	-0.86 (-1.99 to 0.26)	TCAs	1.12 (0.37 to 1.87)	Long-term PDPT individual + AD	2.05 (0.01 to 4.10)
TCAs	-0.54 (-1.18 to 0.10)	Combined (CT/CBT + AD)	0.97 (-0.62 to 2.59)	Combined (CT/CBT + AD)	1.24 (-0.59 to 3.08)
Combined (CT/CBT + AD)	-0.52 (-1.67 to 0.62)	Mirtazapine	0.75 (0.21 to 1.31)	CT/CBT	0.50 (-1.51 to 2.49)
SSRIs	-0.29 (-0.81 to 0.24)	SSRIs	0.56 (-0.04 to 1.15)	IPT	0.51 (-1.20 to 3.04)
Mirtazapine	-0.24 (-0.73 to 0.24)	CT/CBT	-0.06 (-1.68 to 1.60)	TCAs	0.28 (-0.83 to 1.41)
IPT	-0.05 (-1.90 to 1.80)	Pill placebo	reference	SSRIs	0.18 (-0.60 to 0.96)
Pill placebo	reference	IPT	-0.79 (-3.22 to 1.64)	Mirtazapine	0.01 (-1.73 to 1.75)
Behavioural therapies	0.21 (-1.70 to 2.11)	Behavioural therapies	-1.60 (-4.17 to 1.03)	Pill placebo	Reference
Self-help with support	0.59 (-1.51 to 2.67)	Self-help with support	-2.34 (-4.96 to 0.28)	Self-help with support	-1.63 (-6.01 to 2.40)
Counselling	0.69 (-0.92 to 2.34)	Counselling	-2.50 (-4.72 to -0.28)	Waitlist	-2.44 (-6.41 to 1.07)
Short-term PDPT	0.88 (-0.93 to 2.68)	Short-term PDPT	-2.54 (-5.01 to -0.03)		
Self-help without support	0.96 (-0.49 to 2.41)	Exercise	-2.79 (-5.07 to -0.47)		
Exercise	1.04 (-0.54 to 2.60)	Self-help without support	-3.11 (-5.15 to -1.03)		
Attention placebo	1.16 (0.11 to 2.23)	Waitlist	-3.12 (-5.03 to -1.21)		
TAU	1.22 (0.30 to 2.15)	Attention placebo	-3.22 (-5.12 to -1.28)		
Waitlist	1.41 (0.24 to 2.58)	TAU	-3.48 (-5.07 to -1.85)		
Negative values favour classes on the left column		Positive values favour classes on the left column		Positive values favour classes on the left column	
AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; LORs: log-odds ratios; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants					

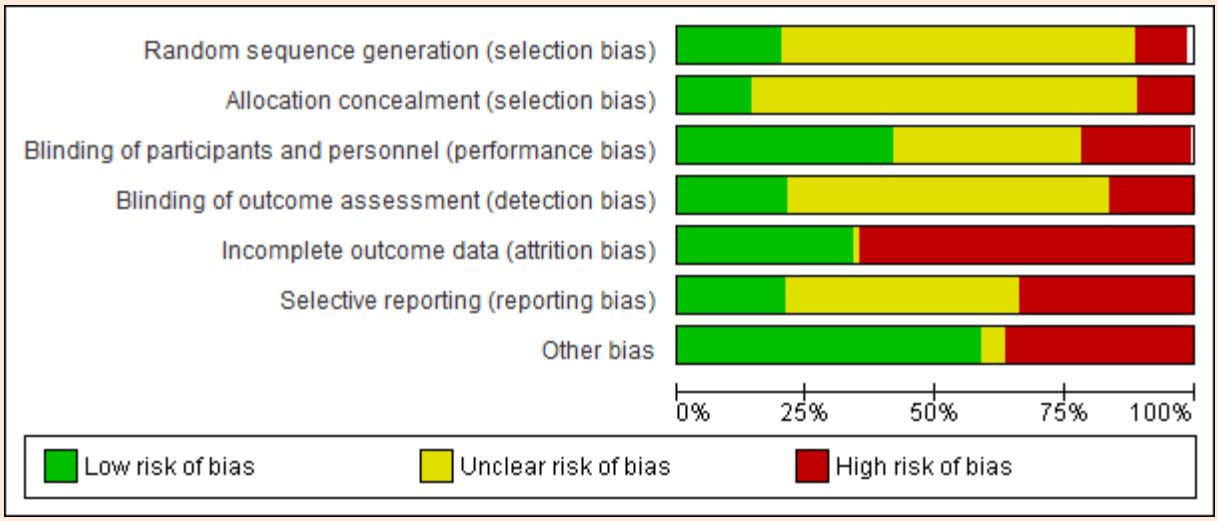
7.5.1.31 Quality of the evidence

The standard GRADE profiles for critical outcomes that have been used to rate the quality of evidence in pairwise meta-analyses conducted for this guideline have not been used for grading the quality in the NMA. This is because GRADE was not developed with network meta-analysis in mind and this is an area of methodological discussion and development. To evaluate the quality of the evidence of the NMAs undertaken to inform this guideline, we report information about the factors that would normally be included in a GRADE profile (i.e. risk of bias, publication bias, imprecision, inconsistency, and indirectness). Study quality and risk of bias were assessed for all studies, irrespective of whether they were included in the network meta-analysis or pairwise comparisons.

Risk of bias

We assessed all included trials for risk of bias (Appendix J3.2). As in the NMA for the less severe network, study reporting was relatively poor and therefore most studies were rated as unclear risk of bias in several domains. Of the studies included in this NMA, 30 were at low risk for sequence generation and 16 of these were at low risk of bias for allocation concealment. Allocation concealment was unclear in 107 trials, and 17 trials were at high risk of bias. Trials of psychological therapies were typically considered at high risk of bias for participant and provider blinding; 61 trials were at low risk of bias for blinding participants and providers, although the rate of side effects may make it difficult to maintain blinding in pharmacological trials as well. Assessor blinding was considered separately for all trials; 30 at low risk of bias, 92 were unclear, and high risk in 23 trials. For incomplete outcome data, 48 trials were at low risk of bias; there was an unclear risk of bias in 3 trials, and 94 trials were at high risk of bias. .Other potential sources of bias were identified in 60 trials. A summary of the risk of bias for these studies is shown in Figure 19.

Figure 19: Risk of bias summary for studies included in the NMA for acute treatment in more severe depression



Model goodness of fit and inconsistency

This section reports only findings of goodness of fit and inconsistency checks for NMA analyses that informed clinical evidence. Detailed findings of goodness of fit and inconsistency checks for all NMA analyses, including those that informed the guideline economic model are reported in the respective sections of Chapter 17.

Update 2017

For the SMD of depressive symptom scores outcome, relative to the size of the intervention effect estimates, moderate between trial heterogeneity was observed for this outcome [τ=0.40 (95% CrI 0.31 to 0.52)]. No meaningful differences were observed in posterior mean residual deviance or between study heterogeneity suggesting that there was no evidence of inconsistency.

For response in those randomised, high between trials heterogeneity was found relative to the size of the intervention effect estimates [τ=0.70 (95% CrI 0.57 to 0.87)]. No meaningful differences were observed in posterior mean residual deviance or between study heterogeneity suggesting that there was no evidence of inconsistency.

For remission in those randomised, high between trials heterogeneity was found relative to the size of the intervention effect estimates, [τ=0.66 (95% CrI 0.43 to 1.05)]. No meaningful differences were observed in posterior mean residual deviance or between study heterogeneity suggesting that there was no evidence of inconsistency.

Detailed comparisons between the relative effects of all pairs of interventions obtained from the consistency (NMA) model and those obtained from the inconsistency (pairwise) model are provided in Appendix W.

Selective outcome reporting and publication bias

The bias adjustment models on SMD of depressive symptom scores that were developed to assess potential bias associated with small study size showed no improvement in fit to the data compared with the unadjusted NMA with the DIC favouring the unadjusted NMA model. However, there was a substantial reduction in the between-study heterogeneity in the bias adjusted model. The mean bias b had a negative median (as expected) and the 95% CrI excluded the possibility of a zero bias although there was large between-study variability in bias [median b=-6.99 (95% CrI -12.77 to -1.19); median standard deviation of b=9.61 (95% CrI 7.16 to 12.74)]. These findings provide moderate evidence of small study bias in comparisons between active and inactive interventions in the SMD outcome.

The SMDs of classes versus pill placebo resulting from the bias adjusted model showed a reduction in relative effect versus pill placebo for most classes (most notably for combined CT/CBT with antidepressant and self-help with support, but reductions in effect were also observed for CT/CBT, combined exercise and antidepressant/CBT - which, nevertheless, remained the most effective class but with high uncertainty around the mean effect -, SSRIs, TCAs and mirtazapine) whereas for counselling, behavioural interventions, short-term psychodynamic psychotherapy, exercise and self-help without support there was an increase in relative effect versus pill placebo. The IPT effect was practically unchanged. Bias-adjusted ranks for classes showed some changes in class ranking. The highest ranked classes (top 6) remained the same but some changes from the base-case analysis were observed in other class rankings and around the uncertainty in rankings. The relative effects of all classes versus pill placebo and versus TAU (posterior mean SMD with 95% CrI) and posterior mean ranks of each class (with 95% CrI) obtained from the bias-adjusted model are provided in Table 54. Classes in the table have been ranked from smallest to largest ranking (with lower rankings suggesting better outcome). The relative effects of every class versus pill placebo obtained from the bias-adjusted model are shown in Figure 20. Table 55 allows comparison of class effects versus pill placebo and class rankings between the base-case results and the bias-adjusted results on the SMD of depressive symptom scores outcome.

For treatment discontinuation, the bias adjusted model showed improved fit to the data compared with the unadjusted NMA, with the DIC favouring the bias-adjusted NMA model, although there was only a small reduction in the between-study heterogeneity when adjusting for bias. The mean bias b had a positive median (as expected) and although the 95% CrI included the possibility of a zero bias, there is a large probability that the bias is indeed positive There was a large variability around the mean bias [median b=0.63 (95% CrI -0.02 to 1.32); standard deviation of b=0.66 (95% CrI 0.16 to 1.19)]. These findings suggest weak

Update 2017

evidence of small study bias in comparisons between active and inactive interventions in the NMA of discontinuation in those randomised.

For response in completers, the bias adjusted model showed some improved fit to the data compared with the unadjusted NMA with a similar DIC between the two models. There was also a small reduction in the between-study heterogeneity in the bias adjusted model. The mean bias had a positive median (as expected) and the 95% CrI excluded the possibility of a zero bias with small variability [median b=1.38 (95% CrI 0.30 to 2.64); standard deviation of b=0.86 (95% CrI 0.03 to 2.08)]. These findings provided evidence of small study bias in this outcome, in comparisons between active and inactive interventions. For this reason, the economic analysis included a probabilistic sensitivity analysis which utilised data on response in completers derived from the bias-adjusted NMA model, to test the impact of the potential small study bias in response in completers outcome on the results of the economic analysis.

Detailed results of all bias models are provided in Appendix N, and Chapter 17, Section 17.8.

Table 54 Results of NMA bias model in people with a new episode of more severe depression. Standardised mean difference (SMD) of depressive symptom scores following adjustment for small study bias: Posterior effects (SMD) of all classes versus pill placebo and TAU and ranking of classes

Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Effect vs TAU (mean, 95% CrI)	Mean rank (95% CrI)
Combined (Exercise + AD/CBT)	41	-0.87 (-2.42 to 0.70) ↓	-1.67 (-3.59 to 0.26) ↓	3.55 (1 to 14)
CT/CBT	391	-0.57 (-1.83 to 0.73) ↓	-1.37 (-2.39 to -0.31) ↓	4.07 (1 to 10)
TCAs	1260	-0.35 (-0.97 to 0.27) ↓	-1.15 (-2.45 to 0.07) ↓	5.00 (1 to 12)
SSRIs	4696	-0.16 (-0.64 to 0.32) ↓	-0.96 (-2.20 to 0.22) ↓	6.44 (2 to 13)
Combined (CT/CBT + AD)	58	-0.18 (-1.64 to 1.23) ↓	-0.98 (-2.79 to 0.76) ↓	6.83 (1 to 16)
Mirtazapine	326	-0.09 (-0.36 to 0.17) ↓	-0.89 (-2.08 to 0.24) ↓	7.08 (3 to 13)
Behavioural therapies	126	-0.08 (-1.96 to 1.84) ↑	-0.87 (-2.42 to 0.68) ↑	7.21 (1 to 16)
IPT	95	0.02 (-2.15 to 2.12) ↓	-0.77 (-2.84 to 1.37) ↓	8.03 (1 to 17)
Pill placebo	2229	reference	-0.80 (-1.97 to 0.28) ↓	8.20 (4 to 14)
Counselling	120	0.26 (-1.44 to 2.04) ↑	-0.54 (-2.02 to 0.87) ↑	9.17 (1 to 17)
Short-term PDPT	115	0.55 (-1.26 to 2.38) ↑	-0.25 (-1.66 to 1.16) ↑	10.78 (2 to 17)
Exercise	50	0.55 (-1.07 to 2.21) ↑	-0.25 (-1.47 to 0.97) ↑	10.85 (2 to 17)
Attention placebo	80	0.63 (-0.49 to 1.82) ↑	-0.16 (-0.77 to 0.43) ↑	11.68 (5 to 16)
Self-help without support	757	0.84 (-0.68 to 2.41) ↑	0.04 (-1.05 to 1.13) ↑	12.70 (4 to 17)
TAU	825	0.80 (-0.28 to 1.97) ↑	reference	12.95 (8 to 16)
Self-help with support	54	1.14 (-1.31 to 3.63) ↓	0.34 (-1.91 to 2.59) ↓	13.04 (2 to 17)
Waitlist	60	1.29 (0.11 to 2.63) ↑	0.49 (-0.23 to 1.16) ↑	15.43 (11 to 17)

Notes:
Negative effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo or TAU)
Arrows next to the class effects indicate whether these have increased (↑) or decreased (↓) compared with the base-case analysis.

Update 2017

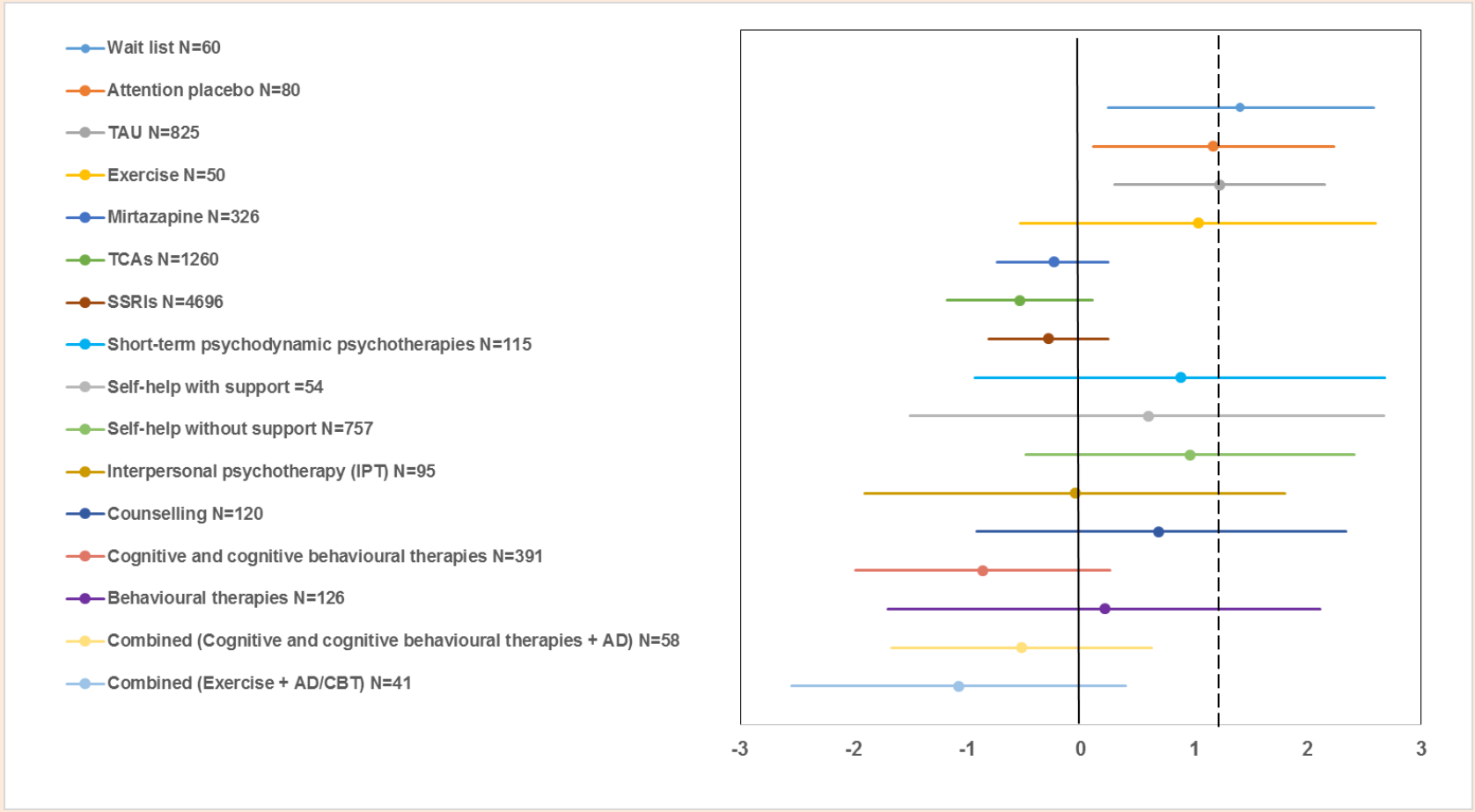
Class	N rand	Effect vs pill placebo (mean, 95% CrI)	Effect vs TAU (mean, 95% CrI)	Mean rank (95% CrI)
AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants				

1

2

Update 2017

Figure 20 Results of NMA bias model in people with a new episode of more severe depression. Standardised mean difference (SMD) of depressive symptom scores of all classes versus pill placebo (N=2229) following adjustment for small study bias [values on the left side of the vertical axis indicate a better effect compared with pill placebo; dotted line indicates TAU effect]



Update 2017

Table 55 Standardised mean difference (SMD) of depressive symptom scores in the NMAs for people with a new episode of more severe depression: comparison between base-case results and results adjusted for small study size bias

Class	N rand	Base-case effect vs pill placebo (mean, 95% CrI)	Base-case mean rank (95% CrI)	Bias-adjusted effect vs pill placebo (mean, 95% CrI)	Bias-adjusted mean rank (95% CrI)
Combined (Exercise + AD/CBT)	41	-1.08 (-2.55 to 0.39)	2.89 (1 to 10)	-0.87 (-2.42 to 0.70) ↓	3.55 (1 to 14)
CT/CBT	391	-0.86 (-1.99 to 0.26)	3.04 (1 to 8)	-0.57 (-1.83 to 0.73) ↓	4.07 (1 to 10)
TCAs	1260	-0.54 (-1.18 to 0.10)	4.33 (1 to 9)	-0.35 (-0.97 to 0.27) ↓	5.00 (1 to 12)
Combined (CT/CBT + AD)	58	-0.52 (-1.67 to 0.62)	4.76 (1 to 11)	-0.18 (-1.64 to 1.23) ↓	6.83 (1 to 16)
SSRIs	4696	-0.29 (-0.81 to 0.24)	5.90 (2 to 11)	-0.16 (-0.64 to 0.32) ↓	6.44 (2 to 13)
Mirtazapine	326	-0.24 (-0.73 to 0.24)	6.25 (2 to 11)	-0.09 (-0.36 to 0.17) ↓	7.08 (3 to 13)
IPT	95	-0.05 (-1.90 to 1.80)	7.41 (1 to 16)	0.02 (-2.15 to 2.12) ↓	8.03 (1 to 17)
Pill placebo	2229	reference	8.17 (5 to 12)	reference	8.20 (4 to 14)
Behavioural therapies	126	0.21 (-1.70 to 2.11)	8.62 (1 to 17)	-0.08 (-1.96 to 1.84) ↑	7.21 (1 to 16)
Self-help with support	54	0.59 (-1.51 to 2.67)	10.43 (1 to 17)	1.14 (-1.31 to 3.63) ↓	13.04 (2 to 17)
Counselling	120	0.69 (-0.92 to 2.34)	11.09 (3 to 17)	0.26 (-1.44 to 2.04) ↑	9.17 (1 to 17)
Short-term PDPT	115	0.88 (-0.93 to 2.68)	11.95 (3 to 17)	0.55 (-1.26 to 2.38) ↑	10.78 (2 to 17)
Self-help without support	757	0.96 (-0.49 to 2.41)	12.50 (6 to 17)	0.84 (-0.68 to 2.41) ↑	12.70 (4 to 17)
Exercise	50	1.04 (-0.54 to 2.60)	12.82 (5 to 17)	0.55 (-1.07 to 2.21) ↑	10.85 (2 to 17)
Attention placebo	80	1.16 (0.11 to 2.23)	13.74 (10 to 17)	0.63 (-0.49 to 1.82)	11.68 (5 to 16)
TAU	825	1.22 (0.30 to 2.15)	14.16 (11 to 17)	0.80 (-0.28 to 1.97) ↑	12.95 (8 to 16)
Waitlist	60	1.41 (0.24 to 2.58)	14.93 (11 to 17)	1.29 (0.11 to 2.63) ↑	15.43 (11 to 17)

Notes

Negative effect values indicate a favourable outcome for classes listed on the first column compared with reference (pill placebo)

Arrows next to the class effects indicate whether these have increased (↑) or decreased (↓) compared with the base-case analysis.

AD: antidepressant; CBT: cognitive behavioural therapy; CT: cognitive therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRIs: selective serotonin uptake inhibitors; TAU: treatment as usual; TCAs: tricyclic antidepressants

Update 2017

Indirectness

In the context of the NMA, indirectness refers to potential differences across the populations, interventions and outcomes of interest, and those included in the relevant studies that informed the NMA.

A key assumption when conducting NMA is that the populations included in all RCTs considered in the NMA are similar. However, it is noted that participants in pharmacological and psychological trials may differ to the extent that some participants find different interventions more or less acceptable in light of their personal circumstances and preferences (so that they might be willing to participate in a pharmacological trial but not a psychological one and vice versa). Similarly, self-help trials may recruit participants who would not seek or accept face-to-face interventions. However, a number of trials included in the NMA have successfully recruited participants who are willing to be randomised to either pharmacological or psychological intervention and to either self-help or face-to-face treatment. The NMAs have assumed that service users are willing to accept any of the interventions included in the analyses; in practice, treatment decisions may be influenced by individual values and goals, and people's preferences for different types of interventions. These factors were taken into account when formulating recommendations.

Interventions of similar type were grouped in classes following GC advice and considered in class models. These models allowed interventions within each class to have similar, but not identical, effects around a class mean effect. Classes and interventions assessed in the NMAs were directly relevant to the classes and interventions of interest.

Outcomes reported in included studies were also the primary outcomes of interest, as agreed by the GC.

7.5.24 Economic evidence

7.5.2.1 Economic literature review

The systematic search of the literature identified 12 UK studies that assessed the cost effectiveness of interventions for adults with a new episode of more severe depression (Benedict et al. 2010; Ekers et al., 2011; Greenhalgh et al. 2005; Hollinghurst et al. 2010; Holman et al. 2011; Horrell et al. 2014; Koeser et al. 2015; Lenox-Smith et al. 2009; Miller et al. 2003; Simon et al. 2006; Wade et al. 2005a and 2005b). Details on the methods used for the systematic search of the economic literature, including inclusion criteria for each review question, are described in Chapter 3. Full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix Q. Completed methodology checklists of the studies are provided in Appendix P. 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 R.

Categorisation of the studies by their population's severity level of depressive symptoms followed the same criteria used for the categorisation of the clinical studies included in the guideline systematic review. All economic studies adopted a NHS perspective, with some studies including personal social service (PSS) costs as well; in addition, some studies reported separate analyses that adopted a societal perspective. NHS and PSS cost elements included, in the vast majority of studies, intervention, primary and community care, staff time (such as GPs, nurses, psychiatrists, psychologists), medication, inpatient and outpatient care and other hospital care. The majority of studies used national unit costs; if a study used different sources for unit costs, this is reported in the text.

Update 2017

7.5.2.1.11 Psychological interventions

Psychoeducation

Horrell and colleagues (2014) evaluated the cost effectiveness of a psychoeducational one-day self-confidence workshop compared with wait list in adults with depression in the UK, alongside a multicentre RCT (Horrell 2014; N=459, completers n=382). The outcome measures of the analysis were the change in BDI-II scores, the number of depression-free days (DFD), calculated based on assumptions around BDI-II scores and the QALY, based on EQ-5D ratings (UK tariff). The duration of the analysis was 12 weeks.

Under a NHS perspective, psychoeducation was found to be overall less costly than wait list. It was reported to be more effective in terms of BDI-II changes and number of DFDs, and produced a similar number of QALYs with wait list. Based on these findings, psychoeducation appeared to be dominant regarding the first two outcomes; regarding QALYs, wait list appeared to be more costly and slightly more effective than psychoeducation with an estimated ICER of £2,472/QALY (2015 prices). The probability of psychoeducation being cost-effective was 0.30, 0.80 and 0.99 at a cost effectiveness threshold of zero, £32 and £74 per BDI-II point improvement, respectively; 0.90 at a cost effectiveness threshold of £15 per DFD gained; and 0.50 at a cost effectiveness threshold of £20,656/QALY, with a maximum probability of 0.56, irrespective of the cost effectiveness threshold per QALY gained. The study is directly applicable to the NICE decision-making context but is characterised by potentially serious limitations mainly due to its short time horizon.

Cognitive behavioural therapy (CBT)

Update 2017

Holman and colleagues (2011) assessed the cost effectiveness of individual CBT versus treatment as usual in older adults with depression in the UK, alongside a RCT (Serfaty 2009; N=204, at endpoint available cost data for n=198, available outcome data for n=167). The study included only primary and community health and personal social care costs; secondary healthcare care costs were not considered. The measure of outcome was the change in BDI-II scores. The time horizon of the analysis was 10 months.

CBT was significantly costlier and more effective than treatment as usual, with an ICER of £137 per additional point reduction in BDI-II (2015 prices). The probability of CBT being cost-effective was 0.90 at a cost effectiveness threshold of £308 per point reduction in BDI-II. Interpretation of these results is difficult as it requires judgements on the value of the unit of outcome. The study is thus only partially applicable to the NICE decision-making context (as no QALYs were used) and is characterised by potentially serious limitations, mainly the omission of secondary healthcare costs from the analysis.

Hollinghurst and colleagues (2010) evaluated the cost effectiveness of individual CBT delivered online using real-time therapist interaction through written messaging versus wait list in people with a new episode of depression in the UK. The economic analysis was undertaken alongside a RCT (Kessler 2009, N=297; BDI data available for n=210; QALYs available for n=165; NHS cost data available for n=137). The outcome measures of the analysis were the change in BDI scores, the percentage of people recovering in each arm, with recovery defined as a BDI score <10, and the QALY, based on EQ-5D ratings (UK tariff). The duration of the analysis was 8 months.

Under a NHS perspective, individual CBT delivered online was significantly more costly than wait list. It was also more effective although the improvement in QALY did not reach statistical significance. Using completers' data, the ICER of CBT with support vs wait list was £4,140 per extra person recovering and £20,150/QALY (2015 prices). The probability of CBT being cost-effective was 0.56 and 0.71 at the NICE lower and upper cost effectiveness thresholds of £20,000 (£23,467 in 2015 prices) and £30,000 (£35,200 in 2015 prices) per

QALY, respectively. After imputation of missing data, the ICER of CBT versus wait list fell at £11,831/QALY, and the probability of CBT being cost-effective rose up to 0.94 and 0.98 at the NICE lower and cost effectiveness thresholds of £20,000 and £30,000/QALY, respectively. The study is directly applicable to the NICE decision-making context and is characterised by potentially serious limitations, mainly the high proportion of missing data.

Behavioural activation

Ekers and colleagues (2011) evaluated the cost effectiveness of behavioural activation delivered over 12 hourly sessions by 2 mental health nurses on post qualification pay bands with no previous formal therapy training for people with a new episode of depression in the UK; therapists received 5-day training and 1 hour clinical supervision fortnightly. The comparator was treatment as usual (TAU), comprising GP care or primary care by mental health workers. The economic analysis was undertaken alongside a RCT (Ekers 2009, N=47; completers n=38). The outcome measures of the analysis were the change in BDI-II scores and the QALY, based on EQ-5D ratings (UK tariff). The duration of the analysis was 3 months. Two alternative scenarios were employed for the cost analysis, based on 2 estimates of workload according to Improving Access to Psychological Therapy (IAPT) service specifications: therapists delivering 65 treatments per year in a depression-specific role (scenario A) or therapists delivering 33 treatments per year treating depression and anxiety (scenario B);

Under a NHS and personal social services perspective, behavioural activation was more costly and more effective than TAU. Using the BDI-II change score as the measure of outcome, the ICER of behavioural activation vs TAU was £10 and £12 per unit change in BDI-II score, for scenarios A and B, respectively (2015 prices). Using the QALY as the measure of outcome and multiple imputation to account for missing data, the ICER of behavioural activation versus TAU was £5,495/QALY (scenario A) or £6,319/QALY (scenario B) in 2015 prices. Following bootstrapping, the probability of CBT being cost-effective was 0.98 and 0.97, for scenarios A and B, respectively, at the NICE lower cost effectiveness threshold of £20,000 (£21,955 in 2015 prices) per QALY. The study is directly applicable to the NICE decision-making context and is characterised by potentially serious limitations, mainly due to its small study side and its short time horizon.

Counselling versus antidepressants

Miller and colleagues (2003) compared the cost effectiveness of counselling (generic psychological therapy comprising 6 weekly 50-minute sessions) versus routinely prescribed antidepressant drugs (mainly dothiepin, fluoxetine or lofepramine) in adults with moderate to severe depression in the UK. The study was conducted alongside a RCT (Bedi 2000; N=103, at 12 months efficacy data for n=81 and resource data for n=103). People refusing randomisation but agreeing to participate in the patient preference trial were given the treatment of their choice (N=220; at 12 months efficacy data for n=163 and resource use data n=215). The study included only depression-related costs. The measure of outcome was a 'global outcome', assessed by a psychiatrist blind to treatment allocation, using the research diagnostic criteria (RDC), the patient's BDI score and GP notes. The outcome was considered good if the person responded to treatment within 8 weeks and then remained well. The outcome measure of the analysis was 12 months.

In the RCT, antidepressants were more costly and more effective than counselling, with an ICER of £483 per extra person with a good global outcome (2015 prices). The probability of counselling being cost-effective was 0.25 and 0.10 at a cost effectiveness threshold of £918 and £3,674 per extra person with a good global outcome, respectively. Sensitivity analysis demonstrated that, assuming missing data reflected good outcomes, the probability of counselling being cost-effective increased at any cost effectiveness threshold; assuming that missing data represented poor outcomes, the probability of counselling being cost-effective

Update 2017

slightly increased for cost effectiveness thresholds lower than £2,755 per good global outcome and decreased for cost effectiveness thresholds higher than £2,755 per good global outcome. In the preference trial, counselling was more costly and more effective than antidepressants with an ICER of £1,675 per extra person with a good global outcome. The study is partially applicable to the NICE decision-making context as it does not use the QALY as the measure of benefit and is characterised by potentially serious limitations, such as the inclusion of depression-related costs only, the use of local unit costs for counsellors, the small numbers of participants randomised as well as included in the preference trial, and the contradictory results between the RCT and the preference trial which did not allow robust conclusions to be drawn.

7.5.2.1.21 Pharmacological interventions

SSRIs versus mirtazapine

Benedict and colleagues (2010) constructed an economic model to evaluate the cost effectiveness of SSRIs and mirtazapine (as well as duloxetine and venlafaxine, which were not part of the decision problem in this review question) in adults with moderate to severe major depression that had a new treatment episode and were treated in primary care in the UK. The duration of the analysis was 48 weeks. Efficacy data were obtained from meta-analyses of RCTs, with randomisation rules possibly being broken. Resource use estimates were based on expert opinion. The outcome measure was the QALY, based on EQ-5D ratings (UK tariff). SSRIs were found to dominate mirtazapine. The results of probabilistic analysis favoured duloxetine, which was not part of the decision problem in this review question. Results were sensitive to the efficacy and utility data. Although the study is directly applicable to the NICE decision-making context, it is characterised by potentially serious limitations, including the methods for meta-analysis and evidence synthesis (selective use of RCTs and synthesis that appears to have potentially broken randomisation) and the fact that it was funded by industry, which may have introduced bias in the analysis.

Update 2017

Fluoxetine versus amitriptyline

Lenox-Smith and colleagues (2009) updated an economic model developed by the same research team (Lenox-Smith et al. 2004) to assess the cost effectiveness of fluoxetine versus amitriptyline (and venlafaxine) in people with depression in the UK. Efficacy data were taken from synthesis of a meta-analysis of trials (fluoxetine versus venlafaxine) and a single trial (amitriptyline versus venlafaxine). The method of synthesis was unclear, but most likely randomisation was broken. Resource use data were elicited from a Delphi panel. The measure of outcome was the QALY, estimated based on the presumed utilities of a depression-free day and a severely depressed day. The time horizon of the analysis was 24 weeks. Fluoxetine was found to dominate amitriptyline, with results being robust to changes in costs but sensitive to the value of the utility gain associated with a depression-free day. The study is partially applicable to the NICE decision-making context (the method of QALY estimation is not consistent with NICE recommendations) and, more importantly, is characterised by very serious limitations, mainly concerning the method of evidence synthesis. Therefore, it has not been considered further when making recommendations.

Escitalopram versus citalopram

Wade and colleagues (2005a and 2005b) undertook model-based economic analysis to assess the cost effectiveness of escitalopram compared with citalopram in adults with major depression (Wade et al. 2005a) and in the subgroup of adults with severe major depression (Wade et al. 2005b). The analyses utilised pooled efficacy data from published RCTs. Resource use data were based on information from a general practice research database, published literature and expert opinion. The measure of outcome was the percentage of

people with remission in each arm of the model, defined as a MADRS score ≤ 12. The time horizon of the analyses was 26 weeks.

In both models, under a NHS perspective, escitalopram dominated citalopram (i.e. it was more effective and less costly). Results were robust to changes in clinical and cost model parameters. In adults with severe depression, escitalopram was dominant in more than 99.8% of the probabilistic analysis iterations. The studies are directly applicable to the NICE decision-making context, as, although the QALY was not used as an outcome, results were straightforward to interpret. However, both studies are characterised by potentially serious limitations, such as the lack of consideration of side effects and their impact on costs and outcomes (study on the whole population of adults with depression), the estimation of resource use based primarily on expert opinion, and the presence of conflicts of interest as both studies were funded by industry.

7.5.2.1.33 Combined psychological and pharmacological interventions

CBT plus antidepressant (fluoxetine) versus antidepressant alone

Simon and colleagues (2006) developed an economic model to assess the cost effectiveness of combination therapy (CBT plus fluoxetine) versus antidepressant (fluoxetine) in adults with moderate or severe depression receiving specialist care in the UK. Efficacy data were derived from a systematic review and meta-analysis of RCTs; resource use data were based on expert opinion and published studies. The outcomes of the analysis were the probability of successful treatment (remission and no relapse over 12 months) with remission defined as HRSD-17 ≤ 6 or HRSD-24 ≤ 8 and the QALY, estimated based on vignettes (descriptions of depression-related health states) valued by service users. The time horizon of the analysis was 15 months.

Using a NHS perspective, combination therapy was found to be more costly and more effective than fluoxetine alone, with an ICER of £5,563 per additional successfully treated person (95% CI £1,920 to £25,099), £19,942/QALY (95% CI £6,583 to £108,901/QALY) for adults with moderate depression, and £7,923/QALY (95% CI £2,606 to 446,358/QALY) for adults with severe depression (2015 prices). Results were sensitive to changes in relative efficacy (in terms of remission and relapse). The authors reported that at the NICE upper cost effectiveness threshold of £30,000/QALY (£41,000/QALY in 2015 price), the probability of combination therapy being cost-effective compared with fluoxetine was 0.88 for adults with moderate depression and 0.97 for adults with severe depression. The study is partially applicable to the NICE decision-making context (as the estimation of QALY was not consistent with NICE recommendations) and is characterised by minor limitations.

Koeser and colleagues (2015) developed an economic model to assess the cost effectiveness of CBT, citalopram and combined therapy of CBT and citalopram in adults with moderate or severe depression receiving specialist care in the UK. Efficacy data for the analysis were derived from systematic screening of a database of RCTs that compared psychological treatments (single or combined) for adults with depression with a control intervention; data were subsequently synthesised using network meta-analysis. Resource use data were based on published estimates of expert opinion and analysis of RCT data. The measure of outcome was the QALY, estimated based on EQ-5D ratings (UK tariff). The time horizon of the analysis was 27 months.

Using a NHS perspective, combination therapy was found to be dominated by CBT, as it was more costly and less effective. CBT was more costly and more effective than citalopram, with an ICER of £20,791/QALY (2015 prices). The probability of each intervention being cost-effective at a cost effectiveness threshold of £26,000/QALY was 0.43 for CBT, 0.37 for citalopram, and 0.20 for combination therapy. Results were sensitive to changes in inclusion criteria for RCTs for acute and follow-up treatment in the systematic review, and the use of SF-6D values (the ICER of CBT versus citalopram reached £33,805/QALY). The study is

Update 2017

directly applicable to the NICE decision-making context and is characterised by minor limitations.

7.5.2.1.43 Physical interventions

ECT

Greenhalgh and colleagues (2006) developed an economic model to assess the cost effectiveness of electroconvulsive therapy (ECT) compared with various pharmacological treatments such as TCAs, SSRIs, SNRIs and lithium augmentation in adults with major depressive disorder who require hospitalisation. The interventions assessed in the analysis were combined in 8 strategies of 3 lines of therapy and maintenance therapy following ECT, which mostly comprised SSRIs. Efficacy data were taken from a systematic literature review of RCTs and published meta-analyses, and further assumptions. Resource use data were based on published literature and expert opinion. The outcome measure was the QALY, estimated based on preferences for vignettes using the McSad health state classification system valued by service users with previous depression in Canada. The time horizon of the analysis was 12 months.

The most effective and cost-effective strategy appeared to be a sequence of ECT – SSRI – lithium augmentation, which had an ICER versus a sequence of SNRI – ECT – lithium augmentation of £9,300/QALY (2015 prices). All other strategies were dominated. Results were modestly sensitive to use of alternative utility values and robust to small changes in costs and suicide rates. The study is partially applicable to the NICE decision-making context as the method of generation of QALYs was not consistent with NICE recommendations and is characterised by potentially serious limitations, including the assumptions made in clinical and cost input parameters.

7.5.2.24 Guideline economic modelling

A decision-analytic model was developed to assess the relative cost effectiveness of pharmacological, psychological and combined interventions for the treatment of a new episode of more severe depression in adults. The objective of economic modelling, the methodology adopted, the results and the conclusions from this economic analysis are described in detail in Chapter 14. This section provides a summary of the methods employed and the results of the economic analysis.

Overview of economic modelling methods

A hybrid decision-analytic model consisting of a decision-tree followed by a three-state Markov model was constructed to evaluate the relative cost effectiveness of a range of pharmacological, psychological and combined interventions for the treatment of a new episode of more severe depression in adults treated in primary care. The time horizon of the analysis was 12 weeks of acute treatment (decision-tree) plus 2 years of follow-up (Markov model). The interventions assessed were determined by the availability of efficacy and acceptability data obtained from the NMAs that were conducted to inform this guideline. Specific interventions were used as exemplars within each class, so that results of interventions can be extrapolated, with some caution, to other interventions of similar resource intensity within their class The following interventions [in brackets the classes they belong to] were assessed:

- pharmacological interventions: sertraline [SSRIs]; mirtazapine [mirtazapine]
- psychological interventions: BA [behavioural therapies]; CBT individual (over 15 sessions) [CT/CBT]; CBT group (under 15 sessions) [CT/CBT]; short term psychodynamic psychotherapy (PDPT) individual [short-term PDPT]; non-directive counselling [Counselling]; cCBT without or with minimal support [self-help without or with mimimal support]

Update 2017

- combined interventions: CBT individual (over 15 sessions) + sertraline [Combined CT/CBT and antidepressant]
- physical interventions: physical exercise programme [exercise]
- clinical management, reflecting GP visits, corresponding to pill placebo RCT arms

The decision-tree component model structure considered the events of discontinuation for any reason and specifically due to intolerable side effects; treatment completion and response reaching remission; treatment completion and response not reaching remission; treatment completion and inadequate or no response. The Markov component model structure considered the states of remission, depressive episode (due to non-remission or relapse), and death. The specification of the Markov component of the model was based on the relapse prevention model developed for this guideline, details of which are provided in Chapter 13.

Efficacy data were derived from the guideline systematic review and NMAs. Baseline parameters (baseline risk of discontinuation, discontinuation due to side effects, response in treatment completers and remission) were estimated based on a review of naturalistic studies. The measure of outcome of the economic analysis was the number of QALYs gained. Utility data were derived from a systematic review of the literature, and were generated using EQ-5D measurements and the UK population tariff. The perspective of the analysis was that of health and personal social care services. Resource use was based on published literature, national statistics and, where evidence was lacking, the GC expert opinion. National UK unit costs were used. The cost year was 2016. Model input parameters were synthesised in a probabilistic analysis. This approach allowed more comprehensive consideration of the uncertainty characterising the input parameters and captured the non-linearity characterising the economic model structure. A number of one-way deterministic sensitivity analyses was also carried out. In addition, a probabilistic sensitivity analysis that used data on response in completers derived from NMAs adjusted for bias resulting from small study size was undertaken.

Results have been expressed in the form of Incremental Cost Effectiveness Ratios (ICERs) following the principles of incremental analysis. Net Monetary Benefits (NMBs) have also been estimated. Incremental mean costs and effects (QALYs) of each intervention versus clinical management (pill placebo) have been presented in the form of cost effectiveness planes. Results of probabilistic analysis have been summarised in the form of cost effectiveness acceptability curves (CEACs), which express the probability of each intervention being cost effective at various cost effectiveness thresholds). Moreover, cost-effectiveness acceptability frontiers (CEAFs) have also been plotted; these show the treatment option with the highest mean NMB over different cost effectiveness thresholds, and the probability that the option with the highest NMB is the most cost-effective among those assessed.

Overview of economic modelling results and conclusions

In people with more severe depression, the combination of CBT individual and sertraline (or another antidepressant) appeared to be the most cost-effective option, with a probability of 0.31 at the NICE lower cost effectiveness threshold of £20,000/QALY. This was followed by CBT group, BA (representing individual behavioural therapies), sertraline (representing SSRIs), physical exercise programme, short term PDPT individual, mirtazapine, counselling, CBT individual, clinical management by GPs (reflecting pill placebo trial arms), and, finally, cCBT without or with minimal support (representing self-help without or with minimal support), which was the least cost-effective option in this population.

Results of the economic analysis were overall robust to different scenarios explored through sensitivity analysis. The relative cost effectiveness of high intensity psychological interventions, alone or combined with antidepressants, improves when these are delivered

Update 2017

by less specialised therapists, such as Band 5 psychological well-being practitioners -PWPs- or Band 6 therapists (instead of Band 7 clinical psychologists) and deteriorates when higher utility values are assumed at baseline, as the scope for HRQoL improvement following successful treatment is more limited. The cost effectiveness of counselling improves if it is delivered in 8 instead of 16 sessions.

Conclusions from the guideline economic analysis refer mainly to people with depression who are treated in primary care for a new depressive episode; however, they may be relevant to people in secondary care as well, given that clinical evidence was derived from a mixture of primary and secondary care settings (however, it needs to be noted that costs utilised in the guideline economic model were mostly relevant to primary care).

Results need to be interpreted with caution due to the limited evidence base characterising some of the interventions assessed in the models, in particular CBT combined with antidepressant and BA, and methodological limitations characterising the NMA on discontinuation for any reason in those randomised, where inconsistency between direct and indirect evidence was identified.

7.5.36 Clinical evidence statements

- Evidence from 41 randomised participants suggests a large but not statistically significant benefit of exercise combined with CBT or an antidepressant relative to pill placebo on depression symptomatology for adults with more severe depression, and this was the highest ranked intervention for clinical efficacy as measured by SMD of depressive symptom scores (mean rank 2.89, 95% CrI 1 to 10).
- Evidence from 391 randomised participants suggests a large but not statistically significant benefit of a cognitive or cognitive behavioural intervention relative to pill placebo on depression symptomatology for adults with more severe depression, and this was the second highest ranked intervention for clinical efficacy as measured by SMD (mean rank 3.04, 95% CrI 1 to 8).
- Evidence from 1260 randomised participants suggests a moderate to large but not statistically significant benefit of a TCA relative to pill placebo on depression symptomatology for adults with more severe depression, and this was the third highest ranked intervention for clinical efficacy as measured by SMD (mean rank 4.33, 95% CrI 1 to 9).
- Evidence from 58 randomised participants suggests a moderate to large but not statistically significant benefit of a cognitive or cognitive behavioural intervention combined with an antidepressant relative to pill placebo on depression symptomatology for adults with more severe depression, and this was the fourth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 4.76, 95% CrI 1 to 11).
- Evidence from 4696 randomised participants suggests a small to moderate, but not statistically significant, benefit of an SSRI relative to pill placebo on depression symptomatology for adults with more severe depression, and this was the fifth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 5.90, 95% CrI 2 to 11).
- Evidence from 326 randomised participants suggests a small to moderate, but not statistically significant, benefit of mirtazapine relative to pill placebo on depression symptomatology for adults with more severe depression, and this was the sixth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 6.25, 95% CrI 2 to 11)).
- Evidence from 95 randomised participants suggests no benefit of IPT relative to pill placebo on depression symptomatology for adults with more severe depression, and this was the seventh highest ranked intervention for clinical efficacy as measured by SMD (mean rank 7.41, 95% CrI 1 to 16), followed by pill placebo, which ranked eighth (mean rank 8.17, 95% CrI 5 to 12).

Update 2017

- Evidence from 126 randomised participants suggests a lower effect of behavioural therapies compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was small to moderate and not statistically significant. Behavioural therapies ranked ninth for clinical efficacy as measured by SMD (mean rank 8.62, 95% CrI 1 to 17).
- Evidence from 54 randomised participants suggests a lower effect of self-help with support compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was moderate to large but not statistically significant. Self-help with support was the tenth highest ranked intervention for clinical efficacy as measured by SMD (mean rank 10.43, 95% CrI 1 to 17).
- Evidence from 120 randomised participants suggests a lower effect of counselling compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was moderate to large but not statistically significant. Counselling was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 11.09, 95% CrI 3 to 17).
- Evidence from 115 randomised participants suggests a lower effect of short-term psychodynamic psychotherapy compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was large but not statistically significant. Short term psychodynamic psychotherapy was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 11.95, 95% CrI 3 to 17).
- Evidence from 757 randomised participants suggests a lower effect of self-help without support compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was large but not statistically significant. Self-help without support was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD (mean rank 12.50, 95% CrI 6 to 17).
- Evidence from 50 randomised participants suggests a lower effect of a physical exercise programme compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was large but not statistically significant. Physical exercise intervention was outside the top-10 highest ranked interventions for clinical efficacy as measured by SMD and ranked below pill placebo (mean rank 12.82, 95% CrI 5 to 17).
- Evidence from 80 randomised participants suggests a lower effect of attention placebo compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was large and statistically significant. Attention-placebo was ranked third from bottom for clinical efficacy as measured by SMD (mean rank 13.74, 95% CrI 10 to 17).
- Evidence from 825 randomised participants suggests a lower effect of treatment as usual compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was large and statistically significant. Treatment as usual was ranked second from bottom for clinical efficacy as measured by SMD (mean rank 14.16, 95% CrI 11 to 17).
- Evidence from 60 randomised participants suggests a lower effect of waitlist compared with pill placebo on depression symptomatology for adults with more severe depression; the difference in effect was large and statistically significant. Waitlist was the bottom ranked intervention for clinical efficacy as measured by SMD (mean rank 14.93, 95% CrI 11 to 17).

Update 2017

7.5.4 Economic evidence statements

7.5.4.1 Psychological interventions

- Evidence from 1 single UK study conducted alongside a RCT (N = 459) suggests that psychoeducation delivered in one day workshop is unlikely to be a cost-effective intervention in people with a new episode of more severe depression. The study is directly applicable to the UK context but is characterised by potentially serious limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N=204) is inconclusive regarding the cost effectiveness of individual CBT in adults with a new episode of more severe depression, as the study did not use the QALY as the measure of outcome, and therefore further judgements are required in order to assess whether the extra unit of benefit gained with CBT is worth its extra cost. The evidence is partially applicable to the NICE decision-making context and is characterised by potentially serious limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N = 297) suggests that computerised CBT delivered online using real-time therapist interaction through written messaging may be a cost-effective intervention in people with a new episode of more severe depression. The study is directly applicable to the UK context but is characterised by potentially serious limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N=47) suggests that behavioural activation delivered by mental health nurses with no previous formal therapy training is likely to be a cost-effective intervention in people with a new episode of more severe depression. The study is directly applicable to the UK context but is characterised by potentially serious limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N= 103) and a preference trial (N= 220) is inconclusive regarding the cost effectiveness of counselling versus antidepressants in adults with a new episode of more severe depression, as the study did not use the QALY as the measure of outcome, and therefore further judgments on cost effectiveness are required. Moreover, results between the RCT and the preference trial were contradictory. The study is partially applicable to the NICE decision-making context and is characterised by potentially serious limitations.

Update 2017

7.5.4.2 Pharmacological interventions

- Evidence from 1 model-based UK study suggests that SSRIs may be more cost-effective than mirtazapine in adults with a new episode of more severe depression. The study is directly applicable to the NICE decision-making context but is characterised by potentially serious limitations.
- Evidence from 1 model-based UK study suggests that fluoxetine may be more cost-effective than amitriptyline in adults with a new episode of more severe depression. However, the study is partially applicable to the NICE decision-making context and is characterised by very serious limitations.
- Evidence from 2 model-based UK studies suggests that escitalopram is more cost-effective than citalopram in adults with a new episode of more severe depression. The evidence is directly applicable to the NICE decision-making context but is characterised by potentially serious limitations.

7.5.4.3 Combined psychological and pharmacological interventions

- Evidence from 1 model-based UK study suggests that combination therapy (CBT and fluoxetine) is likely to be more cost-effective versus pharmacological treatment (fluoxetine) alone in adults with a new episode of more severe depression; evidence is inconclusive of the cost effectiveness of combination therapy in people with moderate-to-severe depression. The evidence is partially applicable to the NICE decision-making context and is characterised by minor limitations.

- Evidence from 1 model-based UK study suggests that CBT is likely to be more cost-effective than combination therapy (CBT and citalopram) in adults with a new episode of more severe depression. The evidence on the cost effectiveness between CBT and pharmacological therapy (citalopram) is inconclusive. The evidence is directly applicable to the NICE decision-making context and is characterised by minor limitations.

7.5.4.4 Physical interventions

- Evidence from 1 model-based UK study suggests that ECT may be cost-effective as part of a sequence of treatments that includes ECT – SSRI – lithium augmentation in adults with major depression that requires hospitalisation. The evidence is partially applicable to the NICE decision-making context and is characterised by potentially serious limitations.

7.5.4.5 Pharmacological, psychological, physical and combined interventions

- Evidence from the guideline economic modelling suggests that the combination of CBT individual and sertraline (or another antidepressant) is likely to be the most cost-effective option for the treatment of new episodes of more severe depression in adults, followed by CBT group, BA (representing individual behavioural therapies), sertraline (representing SSRIs), physical exercise programme, short term PDPT individual, mirtazapine, counselling, CBT individual clinical management by GPs (reflecting pill placebo trial arms), and, finally, cCBT without or with minimal support (representing self-help without or with minimal support). This evidence refers mainly to people treated in primary care for a new depressive episode; however, it may be relevant to people treated in secondary care as well, given that clinical evidence was derived from a mixture of primary and secondary care settings. The economic analysis is directly applicable to the NICE decision-making context and is overall characterised by minor limitations, although the evidence base for some interventions is rather limited, and respective results should therefore be interpreted with caution.

Update 2017

7.6 Subgroup analysis of studies included in the network meta-analysis

This evidence has been synthesised using pair-wise meta-analysis and is relevant to both review questions.

7.6.1.1 Older adults versus younger adults

A comparison of treatments in older and younger adults was believed by the GC to be helpful to inform differential recommendations for this special group. Sufficient data were available to conduct a subgroup analysis of interventions for a new episode of depression in older adults (>60 years of age) compared with younger adults (<60 years of age). No distinction was made between different severity levels for the purpose of the subgroup analysis, however interventions (chosen as they represented some of the main interventions within the NMA) were grouped according to the classes used within the NMA if they were a psychological treatment.

8 RCTs (N=1985) conducted in older adult populations (Bose 2008, Ekkers 2011, Kasper 2005, Laidlaw 2008, Losada 2015, Serfaty 2009, Titov 2015, Tollefson 1993) were compared with RCTs conducted in younger adults across three different comparisons; CBT versus TAU/waitlist, Fluoxetine versus placebo and Escitalopram versus placebo. These comparisons were those with sufficient available data (2 or more studies per comparison) across the interventions most commonly provided and therefore felt to be of greatest interest; SSRIs, CBT, BA and IPT.

An overview of the trials of older adults included in the subgroup analysis can be found in Table 56 and Table 57. Further information about the full NMA included and excluded studies can be found in Appendix J3.1.

Forest plots can be found in Appendix M.

Table 56: Study information table for older adult trials included in the subgroup analysis of CBT versus TAU/waitlist

	CBT versus TAU/waitlist
Total no. of studies (N[1])	5 (530)
Study ID	Ekkers 2011[2] Laidlaw 2008[3] Losada 2015[4] Sefaty 20095 Titov 20156
Country	Netherlands2 UK3,5 Spain4 Australia[6]
Treatment setting	Outpatients[2,4,6] GP[3,5]
Mean age (sd)	COMET + TAU: 71.8 (5.8), TAU: 73.9 (5.7)[2] CBT 74.0(8.4), TAU 74.1(7.6)[3] CBT: 61.5 (12.4), Control: 62.3 (12.9)[4] CBT: 74.4 (7.6), TAU: 72.8 (5.9)[5] Internet CBT: 64.5 (2.6), Control: 66.2 (3.8)[6]
Depression severity	NR
Intervention	CBT; 7x 90 min face-to face group sessions[2], up to 12 x 50min individual face-to-face sessions[5], 8 individual face-to-face weekly sessions[3,4], 8 remote sessions[6]
Comparison	TAU[2,3] Attention placebo[4,5] Waitlist control[6]
Notes: N = total number of participants Ekkers 2011[2,] Laidlaw 2008[3,] Losada 20154, Sefaty 20095, Titov 2015[6]	

Table 57: Study information table for older adult trials included in the subgroup analysis of SSRIs versus other interventions

	Fluoxetine versus placebo	Escitalopram versus placebo
Total no. of studies (N[1])	2 (1,188)	2 (784)
Study ID	Kasper 2005[2] Tollefson 1993[3]	Bose 2008[4] Kasper 2005[2]
Country	Multicentre: BE, CZ, HU, IT, NL, SK, ES, UK[2] NR[3]	NR[4] Multicentre: BE, CZ, HU, IT, NL, SK, ES, UK[2]
Treatment setting	Inpatient[2] NR[3]	NR[4] Inpatient[2]
Mean age (sd)	Fluoxetine: 75 (7), Placebo: 75 (7)[2]	Escitalopram: 68.1 (6.7), Placebo: 68.5 (7.1)[4]

Update 2017

	Fluoxetine versus placebo	Escitalopram versus placebo
	67.7[3]	Escitalopram: 75 (7), Placebo: 75(7)[2]
Depression severity	NR	Moderate-severe[4] NR[2]
Intervention	Fluoxetine; NR[2], 20mg/day[3]	Escitalopram; 10mg/day, increasing to 20mg/day after week 4 if clinically indicated[4], 10mg/day[2]
Comparison	Placebo	Placebo
Notes: N = total number of participants Kasper 2005[2], Tollefson 1993[3], Bose 2008[4]		

There was a large effect of CBT compared with waitlist or TAU in adults over 60 years of age (k=5, n=418), in contrast to younger adults (k=34, n=3,978) in whom CBT had a moderate beneficial effect (SMD=-0.81 [-1.38, -0.25] vs SMD=-0.63 [-0.79, -0.47]). In relation to remission rates, CBT had a stronger but less certain effect (RR 2.90 [0.53, 15.80] vs RR 1.70 [1.33, 2.17]) in older adults (k=2, n=130) than that found in younger adults (k=12, n=1,509). There was clinically important but not statistically significant evidence to suggest that older adults (k=5, n=420) were less likely to discontinue CBT treatment than waitlist or TAU, whilst younger adults (k=39, n=5,644) were more likely to discontinue treatment with CBT than with TAU or waitlist (RR 0.54 [0.24, 1.22] vs RR 1.67 [1.16, 2.40]). For the comparison of CBT versus waitlist or TAU, no data were available for the critical outcome of response.

There was no benefit in terms of remission rates in older adults (k=2, n=928) treated with fluoxetine however in younger adults (k=3, n=776) there was a clinically important but not statistically significant benefit of fluoxetine over placebo (RR 1.01 [0.65, 1.58] vs 1.30 [0.97, 1.73]). Fluoxetine also had no benefit in relation to clinical response in older adults (k=2, n=360) compared with placebo however in younger adults (k=11, n=1,757) there was a clear benefit of fluoxetine (RR 1.16 [0.85, 1.57] vs RR 1.42 [1.21, 1.67]). There was a clinically important but not statistically significant trend (RR 1.53 [0.72, 3.25] vs RR 0.89 [0.76, 1.05]) towards greater discontinuations in older adults (k=2, n=1,014) when treated with fluoxetine but this was not seen in younger adults (k=11, n=2,569).

Escitalopram had no positive effect on remission rates in older adults (k=2, n=509) however there was a statistically significant benefit over placebo (RR 1.10 [0.88, 1.38] vs RR 1.15 [1.01, 1.32]) in younger adults (k=5, n=1,160). The same picture was also seen for response rates with no benefit in older adults (k=2, n=509) but a statistically significant benefit (RR 1.16 [0.93, 1.44] vs 1.21 [1.12, 1.32]) in younger adults (k=8, n=1,918). There were also significantly more discontinuations in the escitalopram group than the placebo group (RR 1.45 [1.03, 2.04] vs RR 1.16 [0.95, 1.41]) in older adults (k=2, n=620) however this pattern was not seen in the younger adult population (k=8, n=2,412).

For the comparisons of fluoxetine and escitalopram versus placebo no data were available for the critical outcomes of efficacy (depressive symptoms) and treatment discontinuations due to side effects.

7.6.1.21 Inpatients versus outpatients

A comparison of treatments in inpatient and outpatient populations was believed by the GC to be helpful in order to examine whether differential recommendations were required for the inpatient population. Sufficient data (2 or more RCTs per comparison) were available to conduct a subgroup analysis of interventions for a new episode of depression in inpatients compared with outpatients for only one comparison; exercise versus attention placebo/TAU.

Update 2017

3 RCTs (N=128) conducted in inpatient populations (Ho 2014, Schuch 2011, Schuch 2015) were compared with RCTs conducted in outpatient populations.

An overview of the trials of inpatients included in the subgroup analysis can be found in Table 58. Further information about the full NMA included and excluded studies can be found in Appendix J3.1.

Forest plots can be found in Appendix M.

Table 58: Study information table for inpatient trials included in the subgroup analysis of exercise versus attention-placebo or treatment as usual

	Exercise versus attention placebo or TAU
Total no. of studies (N[1])	3 (128)
Study ID	Ho 2014[2] Schuch 2011[3] Schuch 2015[4]
Country	Hong Kong[2] Brazil[3,4]
Treatment setting	Inpatient setting
Mean age (sd)	Exercise: 43.6 (13.3), Attention placebo: 48.8 (11.3) Exercise: 42.8 (12.4); Control: 42.5 (13.5) Exercise + TAU: 38.8(11.5),TAU: 41.8(10.4)
Depression severity	NR
Intervention	Exercise
Comparison	Attention placebo[2] No treatment[3] Standard treatment (antidepressants or ECT)[4]
Notes: N = total number of participants Ho 2014[2], Schuch 2011[3], Schuch 2015[4]	

In inpatients (k=2, n=78) there was a weaker but more certain effect (SMD=-0.48 [-0.93, -0.02] vs SMD=-0.58 [-1.36, 0.20]) of exercise on depressive symptoms at treatment endpoint compared with outpatients (k=6, n=236). There was a clinically but not statistically important difference between inpatients (k=3, n=128) and community populations (k=9, n=557) in relation to discontinuation rates when the intervention was exercise (RR 1.48 [0.59, 3.74] vs RR 0.74 [0.49, 1.12]), with inpatients more likely to discontinue and outpatients less likely. Three RCTs provided data on discontinuation rates due to adverse events, however as no events occurred in either group in either the inpatient or community conditions the relative risk ratio was incalculable.

No data were available for the critical outcomes of response or remission.

7.6.2 Clinical evidence statements of sub-group in network meta-analyses

7.6.2.1 Older adults versus younger adults

Cognitive behavioural therapy

Data from five trials (k=5, n=418) showed a large effect of CBT compared with waitlist or TAU in adults over 60 years of age in contrast to younger adults (k=34, n=3,978) in whom CBT had a moderate effect. For remission CBT had a stronger effect in older adults (k=2,

Update 2017

n=130) than that found in younger adults (k=12, n=1,509) and older adults (k=5, n=420) may be less likely to discontinue CBT treatment.

SSRIs

Data from 2 trials in older adults (k=2, n=928) treated with fluoxetine showed no benefit on terms remission however in younger adults (k=3, n=776) there was a clinically important but not statistically significant benefit of fluoxetine over placebo. Fluoxetine also had no benefit in relation to clinical response in older adults (k=2, n=360) compared with placebo however in younger adults (k=11, n=1,757) there was a clear benefit of fluoxetine. There was a non-significant trend (towards greater discontinuations in older adults (k=2, n=1,014) when treated with fluoxetine that was not seen in younger adults (k=11, n=2,569).

Escitalopram had no effect on remission rates in older adults (k=2, n=509) however there was a statistically significant benefit over placebo in younger adults (k=5, n=1,160). The same pattern was seen for response with no benefit in older adults (k=2, n=509) but a significant benefit in younger adults (k=8, n=1,918). There were also more discontinuations with escitalopram than with the placebo group in older adults (k=2, n=620) which was not seen in the younger adult population (k=8, n=2,412).

7.6.2.2 Inpatients versus outpatients

Data from three small trials (N=128) which showed very limited evidence of a small effect exercise on depressive symptoms at treatment endpoint; there was a moderate effect for outpatient populations (k=6, n=236). There were some differences between the two populations in relation to discontinuation, with inpatients more likely to discontinue and outpatients less likely.

7.7 Evidence to recommendations

7.7.1 Relative values of different outcomes

The GC used the results of economic modelling (cost effectiveness) as the main criterion for making recommendations and the NMA results on the SMD of depressive symptom scores outcome (ranking of interventions and relative effects versus pill placebo) as a secondary criterion. Economic modelling was informed by a range of outcomes of the NMAs (discontinuation for any reason, discontinuation due to side effects, response in completers, remission in completers) but not by the SMD outcome. The GC used pill placebo as a benchmark in both the clinical and economic analyses and expressed the view that for an intervention to be recommended, it should show higher cost effectiveness and a better clinical effect compared with pill placebo.

7.7.2 Trade-off between clinical benefits and harms

The GC were predominantly guided by the results of the health economic analysis for those interventions covered by the NMA when drafting the recommendations for people with more severe depression. These recommendations were supported by a review of the relative effectiveness of the interventions against pill placebo.

The GC reviewed the rankings of all interventions and noted that the ranking of the 6 most effective classes of interventions based on the SMD of depressive symptom scores outcome was combined exercise with an antidepressant or cognitive behavioural therapy, cognitive and cognitive behavioural therapy, tricyclic antidepressants, combined cognitive and cognitive behavioural therapy with antidepressants, SSRIs and mirtazapine. For the 3 clinical outcomes assessed (SMD of depressive symptom scores, response in those randomised

Update 2017

and remission in those randomised) the rankings of the classes that ranked in the top six places are summarised below:

- cognitive and cognitive behavioural therapy, combined cognitive and cognitive behavioural therapy with antidepressants, and tricyclic antidepressants were in the top six rankings for all 3 outcomes;
- combined exercise with cognitive behavioural therapy or antidepressants, SSRIs and mirtazapine were in the top six rankings for 2 of the outcomes;
- long-term psychodynamic psychotherapy, long-term psychodynamic psychotherapy with antidepressants and interpersonal psychotherapy were in the top six rankings for 1 outcome.

The GC noted that the inclusion of classes in the top six rankings was affected by data availability. Combined exercise with cognitive behavioural therapy or antidepressants was in the top 6 classes for the outcomes of SMD and response in those randomised, but was not included in the remission in those randomised analysis due to lack of relevant data. Similarly, long-term psychodynamic psychotherapy alone or in combination with antidepressants was in the top six rankings only for the outcome of remission in those randomised; however, this was the only outcome for which long-term psychodynamic psychotherapy data were available. For several classes (behavioural therapies, counselling, short-term psychodynamic psychotherapy, self-help without support, exercise) no data were available on the outcome of remission, and therefore these classes were not included in the respective NMA.

The GC also took into account that there would need to be some flexibility in the treatment options to enable both service user choice and availability of alternative treatment options dependant on past experience of treatment or tolerability problems.

The GC noted the sub-group analysis on older people which showed a relatively larger benefit of CBT for older than younger patients, however, none of the effects on clinical efficacy showed statistically significant subgroup differences and the only outcome that did was for discontinuation which showed that older people were less likely to discontinue CBT than treatment as usual or waitlist (whereas the younger group showed effects in the opposite direction). The GC therefore did not consider it necessary to make differential recommendations for older adults. There were no statistically significant subgroup differences for the inpatient versus outpatient comparison suggesting that differential recommendations were not necessary.

For all severities of depression, the GC agreed that the likely benefits of the recommendations made would be improvements in depression symptoms, remission and response. The potential harms identified were attrition, not taking up of other treatments, issues with acceptability (particularly for drugs which have more side effects) and the possibility of people deteriorating (as data in clinical trials of all treatments estimated this could happen in 7-10% of people). However, the GC agreed that the likely benefits would outweigh the potential harms. In developing the recommendations, the GC also took into account the harm-to-benefit ratio of antidepressants and how the balance of harm and benefit would vary with different severities of depression.

7.7.32 Trade-off between net health benefits and resource use

Existing economic evaluations assessed a limited range of pharmacological, psychological and physical interventions in, mostly, pairwise comparisons, so it was difficult for the GC to draw any robust conclusions on the relative cost effectiveness of the full range of interventions that are available for the treatment of adults with a new episode of more severe depression.

The guideline economic analysis assessed the cost effectiveness of a range of pharmacological, psychological and combined interventions, as well as clinical management

Update 2017

(GP visits, reflected in pill placebo trial arms) as initial treatments for people with a new episode of more severe depression. The interventions included in the economic analysis were dictated by availability of data and were used as exemplars within their class, as for practical reasons it was impossible to model all interventions considered in the guideline NMA. Therefore, the GC noted that results of interventions could be extrapolated, with some caution, to other interventions of similar resource intensity within the same class. It was also noted that due to lack of suitable data, the economic analysis was not able to include IPT and self-help with support and that findings on the SMD outcome for these two interventions were based on limited evidence (N<100 for each).

The GC based the guideline recommendations primarily on the findings of the guideline economic analysis. The ranking of interventions for adults with a new episode of more severe depression, from the most to the least cost-effective was: CBT individual combined with sertraline, CBT group, behavioural activation, sertraline, physical exercise programme, short term psychodynamic psychotherapy, mirtazapine, counselling, CBT individual, clinical management, cCBT without or with minimal support. The GC noted the probabilities of cost effectiveness obtained using a step-wise approach, according to which the most cost-effective intervention is omitted at each step and the probability of the next most cost-effective intervention is re-calculated. The GC also noted that the economic analysis assumed that all individual psychological interventions are delivered by a Band 7 clinical psychologist and that their relative cost effectiveness improved if these were effectively delivered by therapists paid at a lower Band.

The GC took into account the strengths and the limitations of the economic analysis, the robustness of the results under different scenarios explored through sensitivity analysis (including use of data from the NMA bias models), and noted that, with the exception of sertraline and CBT, all other interventions were informed by limited data for some outcomes or borrowed efficacy from a different intervention within their own class or from a different class, depending on availability of appropriate data.

Based on the above considerations, the GC decided to recommend individual CBT in combination with an SSRI (represented by sertraline in the economic analysis) as a first line treatment for more severe depression because it was the most cost-effective intervention in the guideline economic analysis. As a class, combined CBT with an antidepressant also ranked in a high position on the SMD outcome.

The GC recommended group CBT for people with more severe depression who do not want to take medication in combination with psychological therapy, as this was the second most cost-effective intervention and belonged to the class with the second highest ranking on the SMD outcome and the most robust evidence base among psychological interventions.

The GC considered the effectiveness and cost effectiveness of other (individual) psychological interventions: they noted that behavioural activation was the next most cost-effective intervention but showed a negative effect compared with pill placebo on the SMD outcome; short-term psychodynamic psychotherapy and counselling were the next most cost-effective psychological interventions after behavioural activation but also showed a negative effect compared with pill placebo on the SMD outcome; individual CBT ranked in a low place in terms of cost effectiveness, but it was more cost-effective than pill placebo and its class showed the highest effect versus pill placebo;. IPT showed no effect relative to pill placebo on the SMD and it was not included in the economic analysis due to lack of suitable data to inform the economic model. Based on these considerations, the GC decided to recommend individual CBT or behavioural activation for people who do not want to receive combined treatment or group therapy, as the individual psychological intervention with the widest and most robust evidence base and the most cost-effective individual psychological intervention, respectively. The GC considered existing UK evidence according to which behavioural activation can be successfully and cost-effectively delivered by Band 5 therapists in people with more severe depression. The GC noted that the guideline base-case

Update 2017

economic analysis assumed delivery of behavioural activation by Band 7 clinical psychologists and concluded that if behavioural activation is delivered by Band 5 therapists, its relative cost effectiveness will be higher than that estimated by the guideline economic analysis.

The GC considered the relatively high cost effectiveness ranking of short-term psychodynamic psychotherapy, but also its negative effect compared with pill placebo on the SMD outcome. They decided to make a 'consider' recommendation for short-term psychodynamic psychotherapy, alone or in combination with an SSRI or mirtazapine, for people with more severe depression who had had poor response to other recommended interventions (individual CBT in combination with an SSRI, group CBT, individual CBT or behavioural activation) in a previous episode of depression or in those who did not want the other recommended interventions and who would like help for emotional and developmental difficulties in relationships. The GC expressed the view that the effectiveness and cost effectiveness of short-term psychodynamic psychotherapy was likely to be higher in this sub-population compared with the 'general' population with less severe depression that was the focus of the guideline economic analysis.

The GC noted the relative effectiveness and cost effectiveness of SSRIs (sertraline) and mirtazapine, the robust evidence base for both and in particular for SSRIs, and the harm-to-benefit ratio of antidepressants in people with more severe depression and decided to make a recommendation for these drugs for people with more severe depression who do not want psychological treatment.

The GC were concerned that psychological interventions are not always implemented consistently – for example audits have suggested that reduced numbers of sessions are used in practice compared with what is recommended. They therefore agreed it was important to specify the structure of the psychological interventions being recommended to ensure consistency. The recommended structure of all psychological interventions (number and duration of sessions, number of therapists and participants for group interventions) was based on the resource use utilised in the economic analysis, which, in turn, was informed by RCT resource use, modified by the GC expert advice to represent routine clinical practice in the UK, so that recommended structure of psychological interventions represents cost-effective use of available healthcare resources as implemented in routine clinical practice.

Update 2017

7.7.42 Quality of evidence

The GC took into account that evidence for a large number of classes on the SMD outcome was very or moderately limited (exercise N=50; self-help with support N=54; counselling N=120; short-term PDPT N=115; behavioural therapies N=126; IPT N=95; combined exercise with antidepressant/CBT N=41; combined CT/CBT with antidepressant N=58). It was noted that there was no evidence of inconsistency for the SMD outcome. However, there was some evidence of inconsistency for the discontinuation outcome, which informed the economic analysis.

The bias adjustment model on SMD suggested moderate evidence of small study bias in comparisons between active and inactive interventions.

The GC noted that the SMDs of classes versus pill placebo resulting from the bias adjusted model showed a reduction in relative effect versus placebo for most classes (most notably for combined CT/CBT with antidepressant and self-help with support, but reductions in effect were also observed for CT/CBT, combined exercise and antidepressant/CBT - which, nevertheless, remained the most effective class but with high uncertainty around the mean effect -, SSRIs, TCAs and mirtazapine) whereas for counselling, behavioural interventions, short-term psychodynamic psychotherapy, physical exercise programme and self-help without support there was an increase in relative effect versus pill placebo. The IPT effect was practically unchanged. Bias-adjusted ranks for classes showed some changes in class

ranking. The highest ranked classes (top 6) remained the same but some changes from the base-case analysis were observed in other class rankings and around the uncertainty in rankings.

For outcomes used in economic analysis, there was evidence of small study bias in response in completers, in comparisons between active and inactive interventions. However, the GC noted that a sensitivity analysis of the economic model showed economic results to be overall robust to bias adjustment.

Overall, the GC considered that the quality of the evidence, both clinical and economic, was robust enough to allow recommendations to be based on the available evidence.

The GC were also aware that depression is a heterogeneous disorder with a number of different underlying causes and mechanisms. They noted it would be beneficial to identify the mechanism of action of the effective individual psychological treatments for depression to enable the development of better treatments. They therefore recommended further research to fully characterise the nature and range of depressive symptoms experienced by people and relate these to any proposed underlying neuropsychological mechanisms

7.7.56 Other considerations

The GC wanted to compare the findings of the guideline NMAs with those of published reviews and meta-analyses of psychological interventions for people with depression. They noted the different methodology adopted for the guideline NMAs compared with published reviews, which could justify potential differences in results: the guideline NMAs included well-defined populations, without physical comorbidities, who were treated for a new episode of depression; 2 NMAs were conducted separately for people with less severe and people with more severe depression. An important difference between the guideline NMAs and published reviews (including published NMAs) was the inclusion of drug and self-help trials in the analysis. Interventions included in the guideline NMAs were defined and classified differently from other reviews. The guideline NMAs utilised class models, where individual treatment effects are drawn towards a class mean but individual intervention estimates are retained and are more precise. The evidence base used for each NMA analysis was broader than in other reviews, with a combination of continuous (including change from baseline, use of baseline and endpoint mean scores) and dichotomous data being used to inform the SMD and response analyses; a hierarchy of depressive symptom scales was used for this purpose, following GC expert advice.

The GC noted that previous published reviews show superiority of psychological interventions versus control and noticed the difference between published reviews and the guideline NMAs for people with more severe depression. The GC noted the lack of direct, head-to-head comparisons between active psychological interventions in this population. The GC considered the use of pill placebo as the reference treatment and noted that it affected neither the relative effects between classes and interventions nor the rankings of classes and interventions. However, as the pill placebo has a larger effect compared with waitlist and TAU, interventions that appear to be effective compared with waitlist or TAU may not appear to be effective compared with pill placebo, and this may be seen as a difference between previous meta-analyses that have used waitlist or TAU as the reference treatment (comparator) and the guideline NMA that has used pill placebo as the reference treatment. The GC noted that relative effects of interventions versus TAU on the SMD outcome were similar to those observed in published reviews.

The GC discussed the issue of patient choice, with the lay members offering the opinion that many people are happy solely with a choice of either evidence based psychological or pharmacological therapy, with choices between different therapies of the same modality being of less concern. They felt that there would be a subset of patients who would have researched therapies carefully and would have a strong preference, but that this would not

Update 2017

apply to the majority of people. Other issues such as choice of the gender of the therapist, the setting in which interventions were provided and good information on the content of, potential harms or side effects and likely outcomes of an intervention were also considered important.

In developing these recommendations, the GC considered the relative training, experience and salaries of staff providing a range of psychological interventions including counselling, behavioural activation and cognitive behavioural therapies. The GC were aware of the different levels of experience and salary of therapists in some of the trials which form the evidence base. However, the GC took the view that as the majority of high intensity therapists were paid either at AfC Grade 6 or 7 that it was appropriate to use these salaries for the base case economic analysis.

7.8.2 Recommendations

First line treatment for more severe depression

68. Offer individual CBT in combination with an SSRI or mirtazapine as the initial treatment for more severe depression. [new 2017]

69. If a person with more severe depression does not want to take medication, offer:

- group CBT, **or**
- individual CBT or BA if the person does not want group therapy. [new 2017]

70. If a person with more severe depression does not want psychological therapy, offer an SSRI or mirtazapine. [new 2017]

71. Consider short-term psychodynamic psychotherapy, alone or in combination with an SSRI or mirtazapine, for a person with more severe depression who would like help for emotional and developmental difficulties in relationships and:

- has had individual CBT in combination with an SSRI, group CBT, or individual CBT or BA for a previous episode of depression, but this did not work well for them, **or**
- does not want individual CBT in combination with an SSRI, group CBT, or individual CBT or BA. [new 2017]

Update 2017

7.8.3 Research recommendation

1. What are the mechanisms of action of effective psychological interventions for acute episodes of depression in adults?

Statement: A series of experimental studies to identify potential mechanisms associated with current effective treatments for depression should be undertaken and used to inform the development of new treatments. These novel treatments should then be tested in large scale RCTs against current most effective psychological treatments.

Rationale: Depression is a debilitating and highly prevalent condition in adults. Despite significant investment, the most effective and well-established treatments have only modest effects on depressive symptoms, and the majority of treatment is for recurrent depressive episodes. Research is required to identify the mechanism of action of the effective individual psychological treatments for depression, which would allow for the isolation of the most effective components and the development of better treatments. The research will need to be

able to fully characterise the nature and range of depressive symptoms experienced by people and relate these to any proposed underlying neuropsychological mechanisms. The studies will also need to take into account the impact of any moderators of treatment effect. This research is necessary to improve clinical outcomes and quality of life for patients, as well as to reduce the financial burden upon the NHS.

7.9 Pairwise meta-analysis of interventions excluded from the NMA for a new episode of depression

This evidence has been synthesised using pairwise meta-analysis and is relevant to both review questions.

7.9.1.1 Behavioural couples therapy

Five RCTs (N=256) met the eligibility criteria for this review: Beach 1992, Bodenmann 2008, Emanuela-Zurveen 1996, Jacobson 1991, O'Leary 1990.

An overview of the trials included in the meta-analysis can be found in Table 59. Further information about both included and excluded studies can be found in Appendix J4.

Summary of findings can be found in Table 60, Table 61,Table 62 and Table 63. Forest plots and the full GRADE evidence profiles can be found in Appendices M and L.

Across these five RCTs, four comparisons were made: behavioural couples therapy (BCT) versus CBT; BCT versus waitlist control; BCT versus interpersonal psychotherapy (IPT); BCT versus combined BCT and CBT (individual CBT for the depressed wife). No data were available for the critical outcome of response.

Update 2017

Table 59: Study information table for trials included in the meta-analysis of behavioural couples therapy versus waitlist control or active intervention

	BCT versus CBT	BCT versus waitlist control	BCT versus IPT	BCT versus combined BCT and CBT
Total no. of studies (N[1])	5 (256)	2 (88)	1 (60)	1 (72)
Study ID	Beach 1992[2] Bodenmann 2008[3] Emanuels-Zuurveen 1996[4] Jacobson 1991[5] O'Leary 1990[6]	Beach 1992[2] O'Leary 1990[6]	Bodenmann 2008	Jacobson 1991
Country	USA[2,5,6] Germany[3] Netherlands[4]	USA	Germany	USA
Treatment setting	Outpatients	Outpatients	Outpatients	Outpatients
Mean age (SD or range)	Wives 39.14 (28-59), husbands 42.29 (30-69)[2] Depressed patient (by group) CBT: 44.35 (11.31), COCT: 44.35 (10.2). Partner (by group) CBT: 44.95	Wives 39.14 (28-59), husbands 42.29 (30-69)[2] 39.3 (28-59)[6]	Depressed patient (by group) IPT: 47.33 (10.6), COCT: 44.35 (10.2). Partner (by group) IPT: 49.85 (10.26), COCT: 41.85 (10.66)	Wives: 38.5 (8.5), husbands: 40.5 (9.7)

	BCT versus CBT	BCT versus waitlist control	BCT versus IPT	BCT versus combined BCT and CBT
	(11.38), COCT: 41.85 (10.66)[3] 38.2 (8.6)[4] Wives: 38.5 (8.5), husbands: 40.5 (9.7)[5] 39.3 (28-59)[6]			
Depression severity	Milder[3,4,5] More severe[2,6]	More severe	Milder	Milder
Intervention	Behavioural marital therapy: 15-20 face-to-face sessions[2], 16x 1-hour weekly sessions[4] Coping-oriented couples therapy: 10x 2-hour sessions per fortnight[3] Behavioural couples therapy: 20x sessions[5], weekly sessions[6]	Behavioural marital therapy: 15-20 face-to-face sessions[2] Behavioural couples therapy: weekly sessions[6]	Coping-oriented couples therapy: 10x 2-hour sessions per fortnight	Behavioural couples therapy: 20x sessions
Comparison	Individual CBT: 15-20 face-to-face sessions[2], 20x 1-hour weekly sessions[3], 16x 1-hour weekly sessions[4], 20x sessions[5], weekly sessions[6]	Waitlist control	Individual IPT 20x 1-hour weekly sessions	Combined individual CBT (with depressed wife) and behavioural couple therapy, minimum 8x behavioural couple therapy sessions, 6x CBT individual sessions
Notes: N = total number of participants Beach 1992[2], Bodenmann 2008[3], Emanuels-Zuurveen 1996[4], Jacobson 1991[5], O'Leary 1990[6]				

Table 60: Summary of findings table for the comparison of behavioural couples therapy (BCT) and CBT

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	Behavioural couples therapy versus CBT				
Depression symptomatology at endpoint (across severity)		The mean depression symptomatology at endpoint		135 (4 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	SMD 0.03 (-0.49 to 0.54)

Update 2017

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	**Behavioural couples therapy versus CBT**				
BDI/HAMD Follow-up: 10-78 weeks		(across severity) in the intervention groups was **0.03 standard deviations higher** (0.49 lower to 0.54 higher)				
Treatment discontinuation rates (more severe depression) Number of participants discontinuing for any reason Follow-up: mean 15 weeks	**Study population**		**RR 1** (0.25 to 4)	24 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	
	250 per 1000	**250 per 1000** (62 to 1000)				
	Moderate					
	250 per 1000	**250 per 1000** (62 to 1000)				
Depression symptomatology at endpoint (milder depression) BDI/HAMD Follow-up: 16-78 weeks		The mean depression symptomatology at endpoint (milder depression) in the intervention groups was **0.14 standard deviations higher** (0.49 lower to 0.78 higher)		105 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,4]	
Depression symptomatology at endpoint (more severe depression) BDI Follow-up: mean 10 weeks	The mean depression symptomatology at endpoint (more severe depression) in the control groups was **10.87**	The mean depression symptomatology at endpoint (more severe depression) in the intervention groups was **0.34 standard deviations lower** (1.07 lower to 0.38 higher)		30 (1 study)	⊕⊖⊖⊖ **very low**[1,3]	
	Study population					

Update 2017

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	Behavioural couples therapy versus CBT				
Remission BDI<10	**842 per 1000**	**682 per 1000** (480 to 985)	**RR 0.81** (0.57 to 1.17)	38 (1 study)	⊕⊖⊖⊖ **very low**[1,3]	
	Moderate					
Treatment discontinuation rates (across severity) Number of participants discontinuing for any reason Follow-up: 15-78 weeks	**Study population**		**RR 1.97** (0.98 to 3.98)	142 (4 studies)	⊕⊕⊖⊖ **low**[1,3]	
	129 per 1000	**253 per 1000** (126 to 512)				
	Moderate					
	155 per 1000	**305 per 1000** (152 to 617)				
Treatment discontinuation rates (milder depression) Number of participants discontinuing for any reason Follow-up: 16-78 weeks	**Study population**		**RR 2.49** (1.11 to 5.61)	118 (3 studies)	⊕⊕⊖⊖ **low**[1,5]	
	103 per 1000	**258 per 1000** (115 to 580)				
	Moderate					
	143 per 1000	**356 per 1000** (159 to 802)				

Notes:

[1] High or unclear ROB in most domains
[2] I2 <80% but >50%
[3] 95% confidence interval crosses one clinical decision threshold
[4] 95% CI crosses two clinical decision thresholds
[5] Events<300

1

Update 2017

Table 61: Summary of findings table for the comparison of behavioural couples therapy (BCT) and waitlist 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	Behavioural couples therapy versus waitlist control				
Depression symptomatology at endpoint (more severe depression) BDI Follow-up: mean 10 weeks		The mean depression symptomatology at endpoint (more severe depression) in the intervention groups was **12.07 lower** (18.32 to 5.82 lower)		30 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
Treatment discontinuation rates (more severe depression) Number of participants discontinuing for any reason Follow-up: mean 15 weeks	**Study population**		**RR 7** (0.4 to 122.44)	24 (1 study)	⊕⊖⊖⊖ **very low**[1,3]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				

Notes:

[1] High or unclear ROB in most domains
[2] OIS not met (<400 participants)
[3] 95% CI crosses two clinical decision thresholds

Table 62: Summary of findings table for the comparison of behavioural couples therapy (BCT) and interpersonal psychotherapy (IPT)

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	Behavioural couples therapy versus IPT				
Depression symptomatology at endpoint (milder depression) BDI Follow-up: mean 78 weeks		The mean depression symptomatology at endpoint (milder depression) in the intervention groups was **1.56 higher** (5.07 lower to 8.19 higher)		40 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	

Update 2017

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	Behavioural couples therapy versus IPT				
Treatment discontinuation rates (milder depression) Number of participants discontinuing for any reason Follow-up: mean 78 weeks	**Study population**		**RR 1** (0.16 to 6.42)	40 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	
	100 per 1000	**100 per 1000** (16 to 642)				
	Moderate					
	100 per 1000	**100 per 1000** (16 to 642)				

Notes:

[1] High or unclear ROB in most domains
[2] 95% CI crosses one clinical decision threshold
[3] Data not reported for all outcomes
[4] 95% CI crosses two clinical decision thresholds

Table 63: Summary of findings table for the comparison of behavioural couples therapy (BCT) and combined BCT and CBT (with the depressed individual)

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	Behavioural couples therapy versus combined BCT and CBT (individual CBT for the depressed wife)				
Depression symptomatology at endpoint (milder depression) HAMD		The mean depression symptomatology at endpoint (milder depression) in the intervention groups was **4.12 higher** (0.66 lower to 8.9 higher)		40 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
Remission (milder depression) BDI<10	**Study population**		**RR 1.2** (0.74 to 1.94)	40 (1 study)	⊕⊖⊖⊖ **very low**[1,3]	
	571 per 1000	**686 per 1000** (423 to 1000)				
	Moderate					

Update 2017

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	Behavioural couples therapy versus combined BCT and CBT (individual CBT for the depressed wife)				
	571 per 1000	**685 per 1000** (423 to 1000)				
Treatment discontinuation rates (milder depression) Number of participants discontinuing for any reason	**Study population**		**RR 13.36** (0.81 to 218.99)	48 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				

Notes:

[1] High or unclear ROB in most domains
[2] 95% CI crosses one clinical decision threshold
[3] 95% CI crosses two clinical decision thresholds

Update 2017

7.9.1.22 Acupuncture

Seven RCTs (N=1262) met the eligibility criteria for this review and provided data for six comparisons. Two of these RCTs (N=104) compared acupuncture with sham acupuncture (Andreescu 2011; Quah-Smith 2013), two (N=255) compared acupuncture combined with an SSRI with an SSRI-only (Duan 2009; Qu 2013), one (N=75) compared acupuncture with fluoxetine (Sun 2013), one (N=73) compared acupuncture combined with fluoxetine relative to sham acupuncture combined with fluoxetine (Zhang 2013), and one of these RCTs (N=755) had three arms allowing the comparison of acupuncture combined with treatment as usual relative to treatment as usual only], and acupuncture combined with treatment as usual compared with counselling combined with treatment as usual (MacPherson 2013).

An overview of the trials included in the meta-analysis can be found in Table 64 and Table 65. Further information about both included and excluded studies can be found in Appendix J4.

Summary of findings can be found in Table 66, Table 67, Table 68, Table 69, Table 70 and Table 71**Error! Reference source not found.** Forest plots and the full GRADE evidence profiles can be found in Appendices M and L, respectively.

Table 64: Study information table for trials included in the meta-analysis of acupuncture versus sham acupuncture or active intervention

	Acupuncture versus sham acupuncture	Acupuncture + SSRI versus SSRI	Acupuncture versus fluoxetine
Total no. of studies (N[1])	2 (104)	2 (255)	1 (75)
Study ID	Andreescu 2011[2] Quah-Smith 2013[3]	Duan 2009[4] Qu 2013[5]	Sun 2013
Country	US[2] Australia[3]	China	China
Treatment setting	Outpatient	Inpatient[4] Outpatient[5]	Outpatient
Mean age (SD or range)	47.5 (12.7)[2] 38.3 (9.8)[3]	37.5 (10.7)[4] 33.3 (9.7)[5]	42.0 (12.5)
Depression severity	Milder depression	More severe depression	Milder depression
Intervention	Electroacupuncture: 2x30min sessions/week for 12 weeks[2] Laser acupuncture: 2x per week for 4 weeks, then 1x per week for 4 weeks[3]	Electroacupunture + fluoxetine: 6x 30-min sessions/week electroacupunture over 6 weeks + 20mg/day fluoxetine[4] Manual/electroacupunture + paroxetine: 3x 30-min sessions/week acupuncture over 6 weeks + 20mg/day paroxetine[5]	Electroacupuncture: 30 sessions (5x 30-min sessions/week) over 6 weeks
Comparison	Sham electrostimulation: 2x 30-min sessions/week for 6 weeks[2] Placebo acupuncture: 2 sessions/week for 4 weeks, then 1 session/week for 4 weeks[3]	Fluoxetine: 20mg/day for 6 weeks[4] Paroxetine: 20mg/day, initiated at 10mg/day and escalated to 40mg/day if necessary[5]	Fluoxetine: 20mg/day for 6 weeks
Notes: N = total number of participants Andreescu 2011[2], Quah-Smith 2013[3], Duan 2009[4], Qu 2013[5]			

Table 65: Study information table for trials included in the meta-analysis of acupuncture versus sham acupuncture or active intervention (continued)

	Acupuncture + fluoxetine versus sham acupuncture + fluoxetine	Acupuncture + TAU versus TAU	Acupuncture + TAU versus Counselling + TAU
Total no. of studies (N[1])	1 (73)	1 (755)	1 (755)
Study ID	Zhang 2013	MacPherson 2013	MacPherson 2013
Country	China	UK	UK
Treatment setting	Outpatient	Outpatient	Outpatient

Update 2017

	Acupuncture + fluoxetine versus sham acupuncture + fluoxetine	Acupuncture + TAU versus TAU	Acupuncture + TAU versus Counselling + TAU
Mean age (SD or range)	47.2 (9.8)	43.5 (13.4)	43.5 (13.4)
Depression severity	Milder depression	Milder depression	Milder depression
Intervention	Electroacupuncture: 9 sessions (3x 30-min sessions/week) and fluoxetine: 10-40mg/day over 3 weeks	Acupuncture: 12 sessions over 13 weeks, actual mean sessions = 10.3 (3.14) and usual care	Acupuncture: 12 sessions over 13 weeks, actual mean sessions = 10.3 (3.14) and usual care
Comparison	Sham acupuncture: 9 sessions (3x 30-min sessions/week) and fluoxetine: 10-40mg/day over 3 weeks	Usual care (both NHS and private, was available according to need and monitored for all patients)	Counselling (humanistic approach): 12 sessions over 13 weeks, actual mean = 9.0 (3.74) and usual care
Notes: N = total number of participants			

Table 66: Summary of findings table for the comparison of acupuncture versus sham acupuncture

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Acupuncture versus sham acupuncture (95% CI)
Discontinuation due to side effects - Milder symptom severity	107 (2 studies) 8-12 weeks	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 3.1** (0.13 to 73.12)	**Study population**	
				0 per 1000	**-**
				Moderate	
				0 per 1000	**-**
Discontinuation for any reason - Milder symptom severity	104 (2 studies) 8-12 weeks	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.92** (0.24 to 3.55)	**Study population**	
				157 per 1000	**13 fewer per 1000** (from 119 fewer to 400 more)
				Moderate	
				143 per 1000	**11 fewer per 1000** (from 109 fewer to 365 more)
				Study population	

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Acupuncture versus sham acupuncture (95% CI)
Remission – Milder symptom severity HAMD endpoint score of 7 or below	47 (1 study) 8 weeks	⊕⊕⊖⊖ **low**[3,4] due to risk of bias, imprecision	**RR 12.32** (1.76 to 86.26)	**45 per 1000**	**515 more per 1000** (from 35 more to 1000 more)
				Moderate	
				46 per 1000	**521 more per 1000** (from 35 more to 1000 more)
Response – Milder symptom severity HAMD reduction of at least 50% from the baseline score	47 (1 study) 8 weeks	⊕⊕⊖⊖ **low**[3,4] due to risk of bias, imprecision	**RR 3.96** (1.58 to 9.93)	**Study population**	
				182 per 1000	**538 more per 1000** (from 105 more to 1000 more)
				Moderate	
				182 per 1000	**539 more per 1000** (from 106 more to 1000 more)
Depression symptomatology - Milder symptom severity HAMD; endpoint/change score; completer analysis	88 (2 studies) 8-12 weeks	⊕⊖⊖⊖ **very low**[1,5,6] due to risk of bias, inconsistency, imprecision			The mean depression symptomatology - mild/moderate symptom severity in the intervention groups was **2.86 lower** (9.06 lower to 3.34 higher)

Notes:
[1] Randomisation method and method for allocation concealment are not reported
[2] 95% CI crosses line of no effect and two clinical decision thresholds (RR 0.8 and 1.25) and events<300
[3] Allocation sequence not concealed
[4] Events<300
[5] I-squared is over 80%
[6] 95% CI crosses line of no effect and two clinical decision thresholds (SMD -0.5 and 0.5)

Update 2017

Table 67: Summary of findings table for the comparison of acupuncture + SSRI versus SSRI

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Acupuncture + SSRI (fluoxetine/paroxetine) versus SSRI (fluoxetine/paroxetine) (95% CI)
				Study population	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Acupuncture + SSRI (fluoxetine/paroxetine) versus SSRI (fluoxetine/paroxetine) (95% CI)
Discontinuation due to side effects – More severe symptom severity	255 (2 studies) 6 weeks	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.95** (0.25 to 3.71)	**42 per 1000**	**2 fewer per 1000** (from 32 fewer to 114 more)
				Moderate	
				42 per 1000	**2 fewer per 1000** (from 32 fewer to 114 more)
Discontinuation for any reason – More severe symptom severity	255 (2 studies) 6 weeks	⊕⊖⊖⊖ **very low**[2,3] due to risk of bias, imprecision	**RR 0.92** (0.39 to 2.17)	**Study population**	
				84 per 1000	**7 fewer per 1000** (from 51 fewer to 99 more)
				Moderate	
				84 per 1000	**7 fewer per 1000** (from 51 fewer to 98 more)
Remission – More severe symptom severity HAMD endpoint score of 7 or below	157 (1 study) 6 weeks	⊕⊖⊖⊖ **very low**[2,3] due to risk of bias, imprecision	**RR 1.12** (0.61 to 2.06)	**Study population**	
				229 per 1000	**28 more per 1000** (from 89 fewer to 243 more)
				Moderate	
				229 per 1000	**27 more per 1000** (from 89 fewer to 243 more)
Response – More severe symptom severity HAMD reduction of at least 50% from the baseline score	252 (2 studies) 6 weeks	⊕⊖⊖⊖ **very low**[1,4,5] due to risk of bias, inconsistency, imprecision	**RR 1.37** (0.91 to 2.06)	**Study population**	
				453 per 1000	**167 more per 1000** (from 41 fewer to 480 more)
				Moderate	
				453 per 1000	**168 more per 1000** (from 41 fewer to 480 more)
Depression symptomatology – More severe symptom severity HAMD; endpoint/change score; completer analysis	233 (2 studies) 6 weeks	⊕⊕⊖⊖ **low**[1,6] due to risk of bias, imprecision			The mean depression symptomatology - moderate/severe symptom severity in the intervention groups was **0.57 standard deviations lower** (0.84 to 0.29 lower)

Notes:
[1] Randomisation method and method for allocation concealment not reported and no attempt at blinding participants or personnel

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Acupuncture + SSRI (fluoxetine/paroxetine) versus SSRI (fluoxetine/paroxetine) (95% CI)

[2] 95% CI crosses line of no effect and both clinical decision thresholds (RR 0.8 and 1.25) and events<300
[3] No attempt at blinding participants or personnel
[4] I-squared is over 50%
[5] 95% CI crosses both line of no effect and clinical decision threshold (RR 1.25) and events<300
[6] 95% CI crosses clinical decision threshold (SMD -0.5) and N<400

Table 68: Summary of findings table for the comparison of acupuncture versus fluoxetine

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	Acupuncture versus fluoxetine				
Discontinuation due to side effects - Mild/moderate symptom severity Follow-up: mean 6 weeks	No drop-out	No drop-out	Not estimable	75 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
Discontinuation for any reason - Mild/moderate symptom severity Follow-up: mean 6 weeks	**Study population**		**RR 14.78** (0.92 to 28.15) No drop-out in control arm	75 (1 study)	⊕⊖⊖⊖ **very low**[1,3]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				
Response - Mild/moderate symptom severity HAMD reduction of at least 50% from the baseline score Follow-up: mean 6 weeks	**Study population**		**RR 1.25** (0.86 to 1.81)	61 (1 study)	⊕⊕⊖⊖ **low**[1,3]	
	600 per 1000	**750 per 1000** (516 to 1000)				
	Moderate					
	600 per 1000	**750 per 1000** (516 to 1000)				
Depression symptomatology - Mild/moderate symptom severity HAMD; endpoint		The mean depression symptomatology - mild/moderate symptom severity in		61 (1 study)	⊕⊕⊖⊖ **low**[1,4]	

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	Acupuncture versus fluoxetine				
score; completer analysis Follow-up: mean 6 weeks		the intervention groups was **2.45 lower** (4.39 to 0.51 lower)				

Notes:
[1] No attempt at blinding and high risk of attrition bias
[2] Events<300
[3] 95% CI crosses a clinical decision threshold (RR 1.25) and events<300
[4] 95% CI crosses clinical decision threshold (SMD -0.5) and N<400

Table 69: Summary of findings table for the comparison of acupuncture + fluoxetine versus sham acupuncture + fluoxetine

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	Acupuncture + fluoxetine versus sham acupuncture + fluoxetine				
Discontinuation due to side effects - Mild/moderate symptom severity Follow-up: mean 3 weeks	**Study population**		**RR 2.3** (0.48 to 11.11)	73 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	57 per 1000	**131 per 1000** (27 to 635)				
	Moderate					
	57 per 1000	**131 per 1000** (27 to 633)				
Discontinuation for any reason - Mild/moderate symptom severity Follow-up: mean 3 weeks	**Study population**		**RR 1.84** (0.5 to 6.81)	73 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	86 per 1000	**158 per 1000** (43 to 584)				
	Moderate					
	86 per 1000	**158 per 1000** (43 to 586)				
Depression symptomatology - Mild/moderate symptom severity HAMD; change		The mean depression symptomatology - mild/moderate symptom severity in the intervention		70 (1 study)	⊕⊕⊖⊖ **low**[1,3]	

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	Acupuncture + fluoxetine versus sham acupuncture + fluoxetine				
score; ITT analysis Follow-up: mean 3 weeks		groups was **4.68 lower** (7.62 to 1.74 lower)				

Notes:

[1] Method of randomisation not reported and significant difference between groups at baseline in proportion of females (69.4% in intervention relative to 97.1% in control). Allocation concealment method is also not reported. Personnel also non-blind and blinding of outcome assessor not reported
[2] 95% CI crosses line of no effect and both clinical decision thresholds (RR 0.8 and 1.25) and events<300
[3] 95% CI crosses clinical decision threshold (SMD -0.5) and N<400

Table 70: Summary of findings table for the comparison of acupuncture + TAU versus TAU

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	Acupuncture + TAU versus TAU				
Discontinuation due to side effects - Mild/moderate symptom severity Follow-up: mean 13 weeks	**Study population**		**RR 1.17** (0.31 to 4.45)	453 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	20 per 1000	**23 per 1000** (6 to 88)				
	Moderate					
	20 per 1000	**23 per 1000** (6 to 89)				
Discontinuation for any reason - Mild/moderate symptom severity Follow-up: mean 13 weeks	**Study population**		**RR 1.26** (0.79 to 2.01)	453 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	139 per 1000	**175 per 1000** (110 to 280)				
	Moderate					
	139 per 1000	**175 per 1000** (110 to 279)				
Depression symptomatology - Mild/moderate symptom severity		The mean depression symptomatology - mild/moderate		377 (1 study)	⊕⊕⊖⊖ **low**[1,3]	

Update 2017

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	Acupuncture + TAU versus TAU				
PHQ-9; endpoint score; completer analysis Follow-up: mean 13 weeks		symptom severity in the intervention groups was **3.3 lower** (4.67 to 1.93 lower)				

Notes:

[1] No attempts at blinding
[2] 95% CI crosses line of no effect and both clinical decision thresholds (RR 0.8 and 1,25)
[3] 95% CI crosses clinical decision threshold (SMD -0.5) and N<400

Table 71: Summary of findings table for the comparison of acupuncture + TAU versus counselling + TAU

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	Acupuncture + TAU versus Counselling + TAU				
Discontinuation due to side effects - Mild/moderate symptom severity Follow-up: mean 13 weeks	**Study population**		**RR 3.5** (0.73 to 16.71)	604 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	7 per 1000	**23 per 1000** (5 to 111)				
	Moderate					
	7 per 1000	**25 per 1000** (5 to 117)				
Discontinuation for any reason - Mild/moderate symptom severity Follow-up: mean 13 weeks	**Study population**		**RR 0.82** (0.59 to 1.13)	604 (1 study)	⊕⊖⊖⊖ **very low**[1,3]	
	215 per 1000	**176 per 1000** (127 to 243)				
	Moderate					
	215 per 1000	**176 per 1000** (127 to 243)				

Update 2017

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	Acupuncture + TAU versus Counselling + TAU				
Depression symptomatology - Mild/moderate symptom severity PHQ-9; endpoint score; completer analysis Follow-up: mean 13 weeks		The mean depression symptomatology - mild/moderate symptom severity in the intervention groups was **1.5 lower** (2.64 to 0.36 lower)		486 (1 study)	⊕⊕⊕⊖ **moderate**[1]	

Notes:

[1] No attempts at blinding
[2] 95% CI crosses line of no effect and both clinical decision thresholds (RR 0.8 and 1.25)
[3] 95% CI crosses both line of no effect and clinical decision threshold (RR 0.8)

7.9.1.32 Nortriptyline in older adults

Four RCTs (N=313) met eligibility criteria for this review and all four of these RCTs compared nortriptyline with placebo (Georgotas 1986; Katz 1990; Nair 1995; White 1984a).

An overview of the trials included in the meta-analysis can be found in Table 71. Further information about both included and excluded studies can be found in Appendix J4.

Summary of findings can be found in Table 72. Forest plots and the full GRADE evidence profiles can be found in Appendices M and L respectively.

Table 71: Study information table for trials included in the meta-analysis of nortriptyline versus placebo in older adults

	Nortriptyline versus placebo
Total no. of studies (N[1])	4 (313)
Study ID	Georgotas 1986[2] Katz1990[3] Nair1995[4] White1984a[5]
Country	USA Canada, Denmark, UK
Treatment setting	Outpatient[2,5] Residential setting[3] Inpatient and outpatient[4]

Update 2017

	Nortriptyline versus placebo
Mean age in years (SD or range)	Nortriptyline: 64.6 (6.4), placebo: 64.7 (7.6)[2] 84[3] Nortriptyline: median=67, Placebo: median=71[4] 37[5]
Depression severity	NR
Intervention	Nortriptyline: 25mg-125mg/day[2], 25mg titrated as needed[3], 25mg-100mg/day[4], 75-150mg/day[5]
Comparison	Placebo pills
Notes: N = total number of participants	

Table 72: Summary of findings table for 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	Nortriptyline versus placebo				
Depression symptomatology at endpoint HAMD		The mean depression symptomatology at endpoint in the intervention groups was **6.24 lower** (9.17 to 3.3 lower)		109 (2 studies)	⊕⊕⊖⊖ **low**[1,2]	
Depression symptomatology at endpoint - milder depression HAMD	The mean depression symptomatology at endpoint - milder depression in the control groups was **21.2**	The mean depression symptomatology at endpoint - milder depression in the intervention groups was **8.10 lower** (13.17 to 3.03 lower)		23 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
Depression symptomatology at endpoint - more severe HAMD	The mean depression symptomatology at endpoint - more severe in the control groups was **17**	The mean depression symptomatology at endpoint - more severe in the intervention groups was **5.3 lower** (8.89 to 1.71 lower)		86 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
Remission at endpoint - milder depression CGI/HAMD	**91 per 1000**	**584 per 1000** (85 to 1000)	**RR 6.42** (0.93 to 44.16)	23 (1 study)	⊕⊕⊖⊖ **low**[1,3]	

Update 2017

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	Nortriptyline versus placebo				
Remission at endpoint CGI/HAMD	**Study population**		**RR 2.62** (1 to 6.85)	148 (3 studies)	⊕⊖⊖⊖ **very low**[1,3,4]	
	303 per 1000	**793 per 1000** (303 to 1000)				
	Moderate					
	150 per 1000	**393 per 1000** (150 to 1000)				
Remission at endpoint - more severe depression CGI/HAMD	**338 per 1000**	**724 per 1000** (274 to 1000)	**RR 2.14** (0.81 to 5.72)	125 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,4]	
Treatment discontinuations due to side effects - milder depression			**RR 5.58** (0.28 to 110.89)	53 (1 study)	⊕⊖⊖⊖ **very low**[1,5]	
Treatment discontinuation	**Study population**		**RR 1.25** (0.85 to 1.82)	193 (2 studies)	⊕⊕⊖⊖ **low**[1,3]	
	309 per 1000	**386 per 1000** (262 to 561)				
	Moderate					
	333 per 1000	**416 per 1000** (283 to 606)				
Treatment discontinuations due to side effects - more severe depression	**29 per 1000**	**263 per 1000** (35 to 1000)	**RR 9.21** (1.24 to 68.31)	73 (1 study)	⊕⊕⊕⊖ **moderate**[6]	
Treatment discontinuations due to side effects	**Study population**		**RR 7.88** (1.49 to 41.65)	126 (2 studies)	⊕⊕⊖⊖ **low**[1,6]	
	16 per 1000	**125 per 1000** (24 to 661)				
	Moderate					
	14 per 1000	**110 per 1000** (21 to 583)				
Notes:						

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	Nortriptyline versus placebo				

[1] High ROB in one domain and unclear in several others
[2] OIS not met (<400 participants)
[3] 95% CI crosses one clinical decision threshold
[4] I2 >50% but <80%
[5] 95% CI crosses two clinical decision thresholds
[6] OIS not met (<300 events)

7.9.1.42 Omega-3 fatty acids

Five RCTs (N =356) met the eligibility criteria for this review. Two of these RCTs (N=219) compared an omega-3 fatty acid with placebo (Ginty 2015; Mischoulon 2015b) and three of these RCTs (N=137) compared omega-3 fatty acid combined with antidepressant medication to placebo combined with antidepressant medication. In two of these RCTs the antidepressant medication was an SSRI (Gertsik 2012, Jayazeri 2008) and in one of these RCTs the omega-3 fatty acid was combined with any antidepressant/TAU, (Park 2015).

An overview of the trials included in the meta-analysis can be found in Table 73. Further information about both included and excluded studies can be found in Appendix J4.

Summary of findings can be found in Table 74 and Table 75. Forest plots and the full GRADE evidence profiles can be found in Appendices M and L respectively.

Table 73: Study information table for trials included in the meta-analysis of omega-3 fatty acids versus placebo

	Omega-3 fatty acids versus placebo	Omega-3 fatty acids plus SSRIs/antidepressants versus placebo plus SSRIs/antidepressants
Total no. of studies (N[1])	2(219)	3 (137)
Study ID	Ginty 2015[2] Michoulon 2015b[3]	Park 2015[4] Gertsik 2012[5] Jayazeri 2008[6]
Country	US[2,3]	South Korea[4] USA[5] Iran[6]
Treatment setting	Outpatient	NR
Mean age in years (sd or range)	20.2 (1.25)[2] 45.8 (12.5)[3]	Omega-3: 43.5 (3.72), Placebo: 39.41 (3.58)[4] 40.5 (10.2)[5] 34.8 (9.7)[6]
Depression severity	Milder	Milder[4] More severe[5,6]

Update 2017

	Omega-3 fatty acids versus placebo	Omega-3 fatty acids plus SSRIs/antidepressants versus placebo plus SSRIs/antidepressants
Intervention	Long-chain omega-3 polyunsaturated fatty acids (LCPUFAs): pills containing 1000 mg EPA and 400 mg DHA[2] n-3 Poly-unsaturated fatty acids (PUFAs): pills containing 1,140 mg of EPA + 600 mg of DHA, Ropufa 75 n-3 ethyl ester[3]	Eicosapentaenoic acid (EPA) or Docosahexaenoic acid (DHA): 1000mg/d of EPA-enriched mix or 1000mg/d of DHA-enriched mix plus TAU/antidepressant medication (67% SSRI; 33% other AD [NDRI, TCA, SNRI]) [4] Omega-3 fatty acids + citalopram: pills containing 450 mg EPA, 100 mg DHA, and 50 mg other omega-3 fatty acids plus citalopram pills (20-40mg/day)[5] Eicosapentaenoic acid (EPA) + fluoxetine: ethyl-EPA soft gels (1000 mg EPA) + fluoxetine (20mg/day)[6]
Comparison	Placebo pills	Placebo (safflower oil with oleic acid) plus TAU/antidepressant medication (53% SSRI; 47% other AD [NDRI, TCA, SNRI]) [4] Placebo + citalopram: 2 capsules of placebo pills containing olive oil + citalopram (20-40mg/day)[5] Placebo + fluoxetine: placebo soft gels contained 550 mg rapeseed oil + 1x fluoxetine capsule (20mg/day) [6]

Notes:
[1]N=total number of participants
[2]Ginty 2015, [3]Mischoulon 2015b, [4]Park 2015, [5]Gertsik 2012, [6]Jayazeri 2008

1 **Table 74: Summary of findings table for omega-3 fatty acids 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	Omega-3 fatty acids versus placebo				
Remission (milder depression) BDI=>10 or HAMD <=7 at endpoint Follow-up: 3-8 weeks	**Study population**		**RR 1.43** (0.48 to 4.29)	217 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	284 per 1000	**406 per 1000** (136 to 1000)				
	Moderate					
	257 per 1000	**368 per 1000** (123 to 1000)				
Response (milder depression) HAMD reduced by >50% at endpoint Follow-up: mean 8 weeks	**Study population**		**RR 0.92** (0.65 to 1.31)	196 (1 study)	⊕⊖⊖⊖ **very low**[2,3]	
	431 per 1000	**396 per 1000** (280 to 564)				
	Moderate					

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	Omega-3 fatty acids versus placebo				
	431 per 1000	**397 per 1000** (280 to 565)				
Treatment discontinuation (milder depression) Number of participants discontinuing for any reason Follow-up: 3-8 weeks	**Study population**		**RR 0.63** (0.32 to 1.24)	219 (2 studies)	⊕⊕⊖⊖ **low**[3,4]	
	173 per 1000	**109 per 1000** (55 to 215)				
	Moderate					
	142 per 1000	**89 per 1000** (45 to 176)				
Discontinuation due to side effects (milder depression) Number of participants discontinuing due to side effects Follow-up: mean 8 weeks	**Study population**		**RR 1.5** (0.06 to 36.32)	196 (1 study)	⊕⊖⊖⊖ **very low**[2,3]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				

Notes:

[1] I-squared >50%
[2] 95% CI crosses two clinical decision thresholds
[3] Data not reported for all outcomes
[4] 95% CI crosses one clinical decision threshold

Table 75: Summary of findings table for omega-3 fatty acids plus SSRIs/antidepressants versus placebo plus SSRIs/antidepressants

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	Omega-3 fatty acids + SSRI/antidepressants versus placebo + SSRI/antidepressants				
Remission (more severe depression) HAMD <=7 at endpoint	**Study population**		**RR 2.44** (0.88 to 6.82)	40 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	182 per 1000	**444 per 1000** (160 to 1000)				
	Moderate					

Update 2017

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	Omega-3 fatty acids + SSRI/antidepressants versus placebo + SSRI/antidepressants				
Follow-up: mean 8 weeks	**182 per 1000**	**444 per 1000** (160 to 1000)				
Response (more severe depression) HAMD reduced by >50% at endpoint Follow-up: mean 8 weeks	**Study population**		**RR 1.62** (0.94 to 2.8)	32 (1 study)	⊕⊖⊖⊖ **very low**[2,3,4]	
	500 per 1000	**810 per 1000** (470 to 1000)				
	Moderate					
	500 per 1000	**810 per 1000** (470 to 1000)				
Treatment discontinuation (across severity) Number of participants discontinuing for any reason Follow-up: 8-12 weeks	**Study population**		**RR 0.85** (0.44 to 1.63)	117 (3 studies)	⊕⊖⊖⊖ **very low**[1,3,5]	
	271 per 1000	**231 per 1000** (119 to 442)				
	Moderate					
	294 per 1000	**250 per 1000** (129 to 479)				
Treatment discontinuation (milder depression) Number of participants discontinuing for any reason Follow-up: mean 12 weeks	**Study population**		**RR 1.13** (0.42 to 3.03)	35 (1 study)	⊕⊖⊖⊖ **very low**[3,5]	
	294 per 1000	**332 per 1000** (124 to 891)				
	Moderate					
	294 per 1000	**332 per 1000** (123 to 891)				
Treatment discontinuation (more severe depression) Number of participants discontinuing for any reason Follow-up: mean 8 weeks	**Study population**		**RR 0.68** (0.29 to 1.62)	82 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,5]	
	262 per 1000	**178 per 1000** (76 to 424)				
	Moderate					
	259 per 1000	**176 per 1000** (75 to 420)				
Discontinuation due to side effects (more severe	**Study population**		**RR 2** (0.2 to 20.33)	82 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,5]	
	24 per 1000	**48 per 1000** (5 to 484)				

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	Omega-3 fatty acids + SSRI/antidepressants versus placebo + SSRI/antidepressants				
depression) Number of participants discontinuing due to side effects Follow-up: mean 8 weeks	**Moderate**					
	25 per 1000	**50 per 1000** (5 to 508)				

Notes:

[1] High or unclear risk in multiple ROB domains
[2] 95% CI crosses one clinical decision threshold
[3] Data not reported for all outcomes
[4] Unclear risk across multiple ROB domains
[5] 95% CI crosses two clinical decision thresholds

7.9.1.52 Psychosocial interventions (peer support)

Update 2017

Two RCTs (N =251) met the eligibility criteria for this review: Griffiths 2012, Stice 2007. Four comparisons were made across these two RCTs: Peer support group versus waitlist; Peer support (online support group) versus attention-placebo; Peer support group versus CBT group; Peer support group versus self-help (without support).

An overview of the trials included in the meta-analysis can be found in Table 76. Further information about both included and excluded studies can be found in Appendix J4.

Summary of findings can be found in Table 77, Table 78, Table 79 and Table 80. Forest plots and the full GRADE evidence profiles can be found in Appendices M and L respectively.

Table 76: Study information table for trials included in the meta-analysis of peer support versus attention-placebo or active intervention

	Peer support versus waitlist	Peer support (online support group) versus attention-placebo control	Peer support group versus CBT group	Peer support group versus self-help (without support)
Total no. of studies (N[1])	1 (86)	1 (240)	1 (69)	1 (47)
Study ID	Stice 2007	Griffiths 2012	Stice 2007	Stice 2007
Country	US	Australia	US	US
Treatment setting	Outpatient	Outpatient	Outpatient	Outpatient
Mean age in years (SD or range)	18.4 (across all arms including non-extracted arms)	Peer support: 44.4 (12.4); attention control: 44.7 (11.34)	18.4 (across all arms including non-extracted arms)	18.4 (across all arms including non-extracted arms)

	Peer support versus waitlist	Peer support (online support group) versus attention-placebo control	Peer support group versus CBT group	Peer support group versus self-help (without support)
Depression severity	Milder depression	Milder depression	Milder depression	Milder depression
Intervention	Supportive-expressive group intervention: 4x 1-hour weekly sessions	Online peer support (wellbeing board): 2x weekly logins + 4x weekly posts	Supportive-expressive group intervention: 4x 1-hour weekly sessions	Supportive-expressive group intervention: 4x 1-hour weekly sessions
Comparison	, Waitlist	Attention control: online health information and monitoring; 12x weekly modules	CBT group: 4x 1-hour weekly sessions	Cognitive bibliotherapy (without support)
Notes: [1]N=total number of participants				

Table 77: Summary of findings table for peer support versus waitlist for depression

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	Peer support group versus waitlist				
Depression symptoms at endpoint (milder depression) BDI Follow-up: mean 4 weeks		The mean depression symptoms at endpoint (milder depression) in the intervention groups was **7.09 lower** (9.77 to 4.41 lower)		86 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	

Notes:

[1] Unclear allocation concealment and non-blind participants, intervention administrators and outcome assessment
[2] N<400
[3] Data is not reported or cannot be extracted for all outcomes

Update 2017

Table 78: Summary of findings table for peer support versus attention-placebo for depression

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	Peer support (online support group) versus attention control				
Treatment discontinuation (milder depression) Number of participants who discontinued for any reason Follow-up: mean 12 weeks	**Study population**		**RR 3.02** (1.65 to 5.52)	171 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
	134 per 1000	**405 per 1000** (221 to 740)				
	Moderate					
	134 per 1000	**405 per 1000** (221 to 740)				

*The basis for the **assumed risk** (e.g. 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).

CI: Confidence interval; **RR:** Risk ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Notes:

[1] Events<300
[2] Data is not reported or cannot be extracted for all outcomes

Update 2017

Table 79: Summary of findings table for peer support versus CBT group for depression

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	Peer support group versus CBT group				
Depression symptoms at endpoint (milder depression) BDI		The mean depression symptoms at endpoint (milder depression) in the intervention groups was		69 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	

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	Peer support group versus CBT group				
Follow-up: mean 4 weeks		**1.72 lower** (4.8 lower to 1.36 higher)				

Notes:

[1] Unclear allocation concealment and non-blind participants, intervention administrators and outcome assessment
[2] 95% CI crosses one clinical decision threshold
[3] Data is not reported or cannot be extracted for all outcomes

Table 80: Summary of findings table for peer support versus self-help (without support) for depression

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	Peer support group versus self-help (without support)				
Depression symptoms at endpoint (milder depression) BDI Follow-up: mean 4 weeks		The mean depression symptoms at endpoint (milder depression) in the intervention groups was **2.87 lower** (6.53 lower to 0.79 higher)		47 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	

Notes:

[1] Unclear allocation concealment and non-blind participants, intervention administrators and outcome assessment
[2] 95% CI crosses one clinical decision threshold
[3] Data is not reported or cannot be extracted for all outcomes

Update 2017

7.9.25 Clinical evidence statements from pairwise meta-analyses

7.9.2.16 Behavioural couples therapy

- Very low quality evidence from 4 RCTs (N=135) suggests no significant differences between acute first-line treatment with BCT and individual CBT on depression symptomatology at endpoint for adults with either more or less severe depression. Very low quality evidence from one of these RCTs (N=38) also suggests no significant difference between BCT and individual CBT on the rate of remission in adults with less severe depression. However, low quality evidence suggests a trend for a higher rate of discontinuation for adults with either more or less severe depression who were receiving

BCT relative to individual CBT, although this effect just misses statistical significance. In the milder depression subgroup (K=3; N=118) the higher discontinuation in the BCT relative to CBT condition is both clinically important and statistically significant

- Very low quality evidence from 1 RCT (NN=30) suggests a clinically important and statistically significant benefit of acute first-line treatment with BCT relative to a waitlist control condition on depression symptomatology at endpoint in adults with more severe depression. However, very low quality evidence from another single RCT (N=24) suggests a clinically important but not statistically significant harm of BCT relative to waitlist in terms of acceptability (as measured by discontinuation)
- Low to very low quality evidence from 1 RCT (N=40) suggests clinically important but not statistically significant benefits of acute first-line treatment with BCT relative to combined BCT and CBT (for the depressed individual) on depression symptomatology at endpoint and acceptability (as measured by discontinuation) for adults with less severe depression. However, evidence from this same study suggests neither a clinically important nor statistically significant difference between BCT and combined BCT and CBT on the rate of remission.

7.9.2.17 Acupuncture

- Low quality evidence from 1 RCT (N=47) suggests clinically important and statistically significant benefits of acupuncture as an acute first-line treatment, relative to sham acupuncture, on the rate of remission and response in adults with less severe depression. However, very low quality evidence from 2 RCTs (N=88) suggests a clinically important but not statistically significant benefit of acupuncture relative to sham acupuncture on depression symptomatology at endpoint. Very low quality evidence from both of these RCTs (N=107) also suggests a clinically important but not statistically significant harm of acupuncture relative to sham acupuncture with higher discontinuation due to side effects observed in the acupuncture arm (the effect on discontinuation for any reason was not clinically important or statistically significant)
- Low quality evidence from 2 RCTs (N=233) suggests a clinically important and statistically significant benefit of acupuncture combined with an SSRI as acute first-line treatment, relative to an SSRI-only, on depression symptomatology at endpoint for adults with more severe depression. Very low quality evidence from these same 2 RCTs (N=252) also suggests a clinically important benefit of acupuncture in addition to an SSRI relative to an SSRI alone on the rate of response, however, this effect is not statistically significant. Very low quality evidence from one of these RCTs (N=157) suggests neither a clinically important nor a statistically significant effect of acupuncture combined with an SSRI (relative to an SSRI-only) on the rate of remission. Very low quality evidence from both of these RCTs (N=255) suggests neither clinically important nor statistically significant differences, between acupuncture in addition to an SSRI and an SSRI only, on acceptability or tolerability (as measured by discontinuation due to side effects and discontinuation for any reason).
- Low quality evidence from 1 RCT (N=61) suggests a moderate and statistically significant benefit of acupuncture as an acute first-line treatment, relative to fluoxetine, on depression symptomatology at endpoint in adults with less severe depression. However, evidence from the same RCT suggests a clinically important but not statistically significant benefit of acupuncture relative to fluoxetine on the rate of response. There was no discontinuation in this study due to side effects, although there was very low quality evidence for a clinically important but not statistically significant harm of acupuncture relative to fluoxetine with higher discontinuation for any reason
- Low to very low quality evidence from 1 RCT (N=70-73) suggests a moderate to large and statistically significant benefit of acupuncture combined with fluoxetine as an acute first-line treatment, relative to sham acupuncture combined with fluoxetine, on depression symptomatology at endpoint for adults with less severe depression. However, evidence

Update 2017

from both these studies suggests a trend for lower acceptability and tolerability of combined acupuncture and fluoxetine, relative to combined sham acupuncture and fluoxetine, with higher discontinuation due to side effects and due to any reason, although these effects are not statistically significant

- Low to very low quality evidence from 1 RCT (N=377-453) suggests a moderate and statistically significant benefit of acupuncture (in addition to TAU) as an acute first-line treatment, relative to TAU-only, on depression symptomatology for adults with less severe depression. However, this study did find a trend for lower acceptability of acupuncture in addition to TAU relative to TAU-only (as measured by discontinuation due to any reason), although this effect was not statistically significant and evidence from the same study suggests neither a clinically important nor statistically significant difference in terms of tolerability (as measured by discontinuation due to side effects)
- Moderate quality evidence from 1 RCT (N=486) suggests a small but statistically significant benefit of acupuncture (in addition to TAU) as an acute first-line treatment, relative to counselling (in addition to TAU), on depression symptomatology at endpoint for adults with less severe depression. However, very low quality evidence from the same RCT (N=604) suggests a clinically important but not statistically significant harm of acupuncture relative to counselling in terms of tolerability (as measured by discontinuation due to side effects), although the absolute numbers are small and no difference was found between acupuncture and counselling in terms of discontinuation for any reason

7.9.2.3 Nortriptyline in older adults

- Low to very low quality evidence from 2-3 RCTs (N=109-148) suggests a large and statistically significant benefit of nortriptyline as an acute first-line treatment, relative to placebo, on depression symptomatology at endpoint in older adults with either less or more severe depression, and a clinically important benefit that just misses statistical significance on the rate of remission. However, low quality evidence from 2 of these RCTs (N=126-193) suggests a clinically important and statistically significant harm of nortriptyline relative to placebo in terms of tolerability (as measured by discontinuation due to side effects) and a clinically important but not statistically significant harm in terms of acceptability (as measured by discontinuation for any reason)

Update 2017

7.9.2.4 Omega-3 fatty acids

- Low to very low quality evidence from 2 RCTs (N=217-219) suggests clinically important but not statistically significant benefits of an omega-3 fatty acid as an acute first-line treatment, relative to placebo, on the rate of remission and acceptability (as measured by discontinuation for any reason) in adults with less severe depression.. However, very low quality evidence from one of these RCTs (N=196) suggests neither a clinically important nor statistically significant effect of an omega-3 fatty acid on the rate of response and evidence from the same study suggests a clinically important but not statistically significant harm associated with an omega-3 fatty acid in terms of tolerability (as measured by discontinuation due to side effects)
- Very low quality evidence from single-study analyses (N=32-40) suggests a clinically important but not statistically significant benefit of omega-3 fatty acid supplementation of SSRI treatment as an acute first-line treatment, compared with placebo augmentation, on the rate of remission and the rate of response in adults with more severe depression. Very low quality evidence from 3 RCTs (N=117) suggests neither a clinically important nor statistically significant difference between omega-3 supplementation and placebo supplementation (of antidepressant medication) on acceptability (as measured by discontinuation for any reason) for adults with either less severe or more severe depression. However, very low quality evidence from 2 of these RCTs (N=82) suggests a clinically important, but not statistically significant, harm of omega-3 supplementation of SSRIs on tolerability (as measured by discontinuation due to side effects) in adults with more severe depression

Psychosocial interventions (peer support)

- Very low quality evidence from 1 RCT (N=86) suggests a large and statistically significant benefit of a peer support group as an acute first-line treatment, relative to waitlist, on depression symptomatology at endpoint for adults with less severe depression
- Low quality evidence from 1 RCT (N=171) suggests a clinically important and statistically significant harm of an online peer support group as an acute first-line treatment, relative to an attention-placebo control (online health information and monitoring), in terms of acceptability (as measured by discontinuation for any reason) for adults with less severe depression
- Very low quality evidence from 1 RCT (N=69) suggests neither a clinically important nor statistically significant difference between a peer support group and a CBT group intervention, as acute first-line treatment, on depression symptomatology at endpoint for adults with less severe depression
- Very low quality evidence from 1 RCT (N=47) suggests neither a clinically important nor statistically significant difference between a peer support group and self-help (without support), as acute first-line treatment, on depression symptomatology at endpoint for adults with less severe depression

Evidence to recommendations

Relative values of different outcomes

Depression symptomology, remission and response were identified as critical outcomes for the pairwise comparisons. Important (but not critical) outcomes were discontinuation due to side effects and discontinuation due to any reason (including side effects).

Trade-off between clinical benefits and harms

The GC agreed that clinical benefits from the interventions examined through pairwise meta-analysis would be improved clinical outcomes, as evidenced by increased remission and response and decreased symptoms. They agreed that behavioural couples therapy, amongst the interventions examined here, appeared to provide this. The potential clinical harms would be higher discontinuation rates, increase in relationship difficulties or a lack of acceptability of the intervention.

Trade-off between net health benefits and resource use

The GC noted that there was no available economic evidence on behavioural couples therapy. However, after reviewing the clinical evidence for this intervention and comparing the effects and related resource use with other psychological interventions that were shown to be cost-effective in the economic analyses (such as CBT or behavioural activation), they decided to make a ‘consider’ recommendation for behavioural couples therapy for people with depression who have a relationship problem if the problem might be related to their depression or if involving their partner may help them with their depression. The GC expressed the view that such a recommendation would have modest resource implications as it affects only those people where relationship problems are contributing to the depression and not everyone in this situation will seek treatment.

Quality of evidence

The GC noted that very low to low quality evidence had been found for acupuncture, nortriptyline in older adults and omega-3 fatty acids. For acupuncture, there was evidence of a statistically significant effect of acupuncture on depressive symptoms compared with SSRIs and higher rates of remission and response in those with less severe depression

Update 2017

when compared with sham acupuncture. There was no statistically significant difference in discontinuation. There was no statistically significant increase in response rates and no difference in remission rates of acupuncture in combination with an SSRI (compared to an SSRI alone) in those with more severe depression. As blinding of provider is typically not possible in these studies, the GC were unsure of the possible impact of this on the findings. They also queried whether the context of the study (4 of the studies were conducted in China) may have impacted upon the apparent efficacy of the intervention. Although they agreed that these were potentially promising results they did not feel able to make recommendations on the basis of the available evidence as they had concerns about the generalisability of the intervention.

For nortriptyline, the evidence suggests nortriptyline is more effective than placebo on depression symptomatology at endpoint in older adults with either less or more severe depression, and may be associated with an increased rate of remission in older people with depression (although this effect was not statistically significant). However, the evidence was from a small number of studies in which higher rates of discontinuation were also seen. For omega-3 fatty acids the evidence showed no statistically significant benefit on remission, response or discontinuation compared with placebo.

The GC noted the low quality of the evidence for acupuncture, nortriptyline and omega-3 fatty acids and the fact that there was a lot of uncertainty over the effectiveness of these interventions. They therefore agreed not to make any recommendations for these interventions.

The GC also noted that very low quality evidence had been found on behavioural couples therapy but with less uncertainty for the other interventions and the GC also had confidence in the generalisability of the findings. Although the evidence was limited it did suggest that behavioural couples therapy may be as effective as individual CBT on depression symptoms at endpoint for adults with less or more severe depression, and is better than a waitlist control condition for depression symptoms at endpoint in adults with more severe depression. The GC were also aware that relationship difficulties are associated both with a poorer response to initial treatment and an increased likelihood of relapse after successful treatment and this further supported their view that a recommendation should be made for behavioural couples therapy

The GC noted that the evidence on peer support was limited and of very low quality. There was single-study evidence for benefits of a peer support group relative to waitlist on depression symptoms at endpoint for adults with less severe depression. However, evidence from another study suggested a higher rate of treatment discontinuation in an online peer support intervention compared with attention-placebo control (online health information) and no differences were found between a peer support group and a CBT group or self-help (without support) intervention on depression symptoms at endpoint for adults with less severe depression. Given this the GC agreed not to make any recommendations for clinical practice. However, they were aware that peer support is a popular intervention and its use is currently being encouraged so they agreed to recommend further research in this area in order to get more data in future that might enable a recommendation for clinical practice to be made.

7.9.3.5 Other considerations

The GC were concerned that psychological interventions are not always implemented consistently – for example audits have suggested that reduced numbers of sessions are used in practice compared with what is recommended. They therefore agreed it was important to specify the structure of the behavioural couples therapy being recommended to ensure consistency in the delivery of this intervention. The recommended structure was based on the manuals that were used in the clinical trials of behavioural couples therapy.

Update 2017

7.9.4 Recommendations

Behavioural couples therapy for depression

72. Consider behavioural couples therapy for a person with more severe and less severe depression who has problems in the relationship with their partner if:

- the relationship problem(s) could be contributing to their depression, **or**
- involving their partner may help in the treatment of their depression. [new 2017]

73. Ensure behavioural couples therapy for people with depression:

- follows the behavioural principles for couples therapy
- provides 15–20 sessions over 5–6 months. [2017]

7.9.5 Research recommendation

2. Is peer support an effective and cost effective intervention in improving outcomes, including symptoms, personal functioning and quality of life in adults as a stand-alone intervention in people with less severe depression and as an adjunct to other evidence based interventions in more severe depression?

Update 2017

Statement: A series of randomised controlled trials should be conducted to assess the effectiveness of different models of peer support which examine the effectiveness and cost effectiveness of peer support for different severities of depression alone or in combination with evidence-based interventions for the treatment of depression. The studies should report on depressive symptoms, personal functioning and quality of life and any adverse events. They should have a follow-up period of at least 12 months.

Rationale: Not all people with depression respond well to first-line treatments and for some people the absence of good social support systems may account for the limited response to first line interventions. A number of models for the provision of peer support have been developed in mental health which aim to provide direct personal support and help with establishing and maintaining supportive social networks. Peer support is provided by people who themselves have personal experience of a mental health problem. However, to date few studies have established and tested peer support models for people with depression. Peer support models, including both individual and group interventions, should be tested in a series of randomised controlled trials which examine the effectiveness of peer support for different severities of depression alone or in combination with evidence-based interventions for the treatment of depression.

7.10 St John's wort

7.10.1 Studies considered[fg]

Forty studies were found in a search of electronic databases, with 19 being included and 21 being excluded by the GDG.

f Details of standard search strings used in all searches are in Appendix H. Information about each study along with an assessment of methodological quality is in Appendix J11, which also contains a list of excluded studies with reasons for exclusions.

g Study IDs in title case refer to studies included in the 2004 guideline. References for these studies are in Appendix U.

Ten studies were available for a comparison with placebo (Davidson02, Hansgen1996, Kalb2001, Laakmann98, Lecrubier02, Philipp99, Schrader98, Shelton2001, Volz2000, Witte1995); four studies for a comparison with TCAs (Bergmann93, Philipp99, Wheatley97, Woelk2000); one for a comparison with TCA-related antidepressants (Harrer94); and six studies for a comparison with SSRIs (Behnke2002, Brenner00, Davidson02, Harrer99, Schrader00, VanGurp02)[h]. Data from up to 1520 participants were available from studies comparing St John's wort with placebo, and data from up to 1629 participants were available from comparison with antidepressants.

All included studies were published between 1993 and 2002 and were between 4 and 12 weeks' long (mean = 6.47 weeks). In 16 studies participants were described as outpatients and in the other three it was either not clear from where participants were sourced or they were from mixed sources. In one study (Harrer99), all participants were aged 60 years and over. All participants had either moderate or severe depression. It is very difficult to assess the exact content of the preparation of St John's wort used in included studies so no study was excluded on grounds of inadequate dose.

Included studies described the following range of preparations:

- 2 X 150 mg (300 mg) at 0.450 to 0.495 mg total hypericin per tablet
- 900 mg LI 160
- 4 X 200 mg (800 mg) LoHyp-57: drug extract ratio 5–7:1
- 3 X 300 mg (900 mg) WS5572: drug extract ratio 2.5–5:1, 5% hyperforin
- 3 X 300 mg (900 mg) WS5573: 0.5% hyperforin
- 3 X 300 mg (900 mg) WS5570: 0.12 to 0.28% hypericin
- 3 X 350 mg (1050 mg) STEI 300: 0.2 to 0.3% hypericin, 2 to 3% hyperforin
- 2 X 200 mg (500 mg) ZE117: 0.5 mg hypericin
- 3 to 6 X 300 mg (900 mg to 1800 mg) at 0.3% hypericum
- 3 X 300 mg (900 mg) LI 160 = 720 to 960 mcg hypericin
- 2 X 250 mg (500 mg) ZE117: 0.2% hypericin
- 900 mg to 1500 mg LI 160: standardised to 0.12 to 0.28% hypericin
- 4 X 125 mg (500 mg) Neuroplant
- 200–240 mg Psychotonin forte
- 3 X 30 drops Psychotonin (500 mg)
- 3 X 30 drops Hyperforat: 0.6 mg hypericin.

In addition, six studies with low doses of standard antidepressants were also included.

7.10.2.4 Clinical evidence statements for St John's wort compared with placebo[i]

7.10.2.5 Effect of treatment on efficacy outcomes

There is some evidence suggesting that there is a clinically important difference favouring St John's wort over placebo on increasing the likelihood of achieving a 50% reduction in symptoms of depression as measured by the HRSD in:

- the dataset as a whole (K = 6139; N = 995; RR = 0.79; 95% CI, 0.71 to 0.88)
- moderate depression (K = 1; N = 162; RR = 0.64; 95% CI, 0.51 to 0.79)

h 137Davidson02 and Philipp99 are 3-arm trials.
i The forest plots can be found in Appendix L

- severe depression (K = 5j; N = 898; RR = 0.81; 95% CI, 0.72 to 0.9).

There is insufficient evidence to determine if there is a clinically important difference between St John's wort and placebo on increasing the likelihood of achieving remission by the end of treatment as measured by the HRSD (K = 3; N = 804; Random effects RR = 0.80; 95% CI, 0.53 to 1.22).

There is evidence suggesting that there is a statistically significant difference favouring St John's wort over placebo on reducing symptoms of depression by the end of treatment as measured by the HRSD, but the size of this difference is unlikely to be of clinical importance in:

- the dataset as a whole (K = 6k; N = 1031; SMD = -0.35; 95% CI, -0.47 to -0.22)
- severe depression (K = 5l; N = 891; SMD = -0.34; 95% CI, -0.47 to -0.2).

However, in moderate depression there is some evidence suggesting that there is a clinically important difference favouring St John's wort over placebo on reducing symptoms of depression by the end of treatment as measured by the HRSD (K = 2; N = 299; Random effects SMD = -0.71; 95% CI, -1.28 to -0.13).

7.10.2.2 Acceptability and tolerability of treatment

There is evidence suggesting that there is no clinically important difference between St John's wort and placebo on reducing the likelihood of patients leaving treatment early for any reason (K = 8; N = 1472; RR = 0.96; 95% CI, 0.74 to 1.25).

There is insufficient evidence to determine if there is a clinically important difference between St John's wort and placebo on reducing the likelihood of patients leaving treatment early due to adverse effects (K = 5; N = 1127; RR = 0.88; 95% CI, 0.32 to 2.41).

There is evidence suggesting that there is no clinically important difference between St John's wort and placebo on reducing the likelihood of patients reporting adverse effects (K = 7; N = 1106; RR = 0.89; 95% CI, 0.72 to 1.1).

7.10.3 Clinical evidence statements for St John's wort compared with antidepressantsm

7.10.3.1 Effect of treatment on efficacy outcomes

There is evidence suggesting that there is no clinically important difference between St John's wort and antidepressants on:

- increasing the likelihood of achieving a 50% reduction in symptoms of depression as measured by the HRSD (K = 10; N = 1612; Random effects RR = 1.03; 95% CI, 0.87 to 1.22)
- increasing the likelihood of achieving remission by the end of treatment as measured by the HRSD (K = 1; N = 224; RR = 1.01; 95% CI, 0.87 to 1.17)
- reducing symptoms of depression by the end of treatment as measured by the HRSD (K = 9; N = 1168; SMD = -0.02; 95% CI, -0.13 to 0.1).

A sub-analysis by severity found no difference in these results except for response rates in those with moderate depression:

j Two studies (Davidson02, Hangsen1996) were removed from the meta-analysis to remove heterogeneity from the dataset.
k Three studies (Davidson02, Hangsen1996, Schrader98) were taken out of the meta-analysis to remove heterogeneity from the dataset.
l Ibid.
m The forest plots can be found in Appendix L

In moderate depression there is some evidence suggesting that there is a clinically important difference favouring St John's wort over antidepressants on increasing the likelihood of achieving a 50% reduction in symptoms of depression as measured by the HRSD (K = 3; N = 481; RR = 0.77; 95% CI, 0.62 to 0.95).

Sub-analyses by antidepressant class and by antidepressant dose (therapeutic versus low dose) found similar results.

A sub-analysis combining severity and antidepressant dose also found similar results apart from for response rates in severe depression:

In severe depression there is some evidence suggesting that there is a clinically important difference favouring low-dose antidepressants over St John's wort on increasing the likelihood of achieving a 50% reduction in symptoms of depression as measured by the HRSD (K = 4; N = 521; RR = 1.2; 95% CI, 1 to 1.44).

7.10.3.2 Acceptability and tolerability of treatment

With regard to reducing the likelihood of patients leaving treatment early for any reason, there is insufficient evidence to determine a difference between St John's wort and either all antidepressants or low-dose antidepressants. However, there is some evidence suggesting that there is a clinically important difference favouring St John's wort over antidepressants given at therapeutic doses (K = 5; N = 1011; RR = 0.69; 95% CI, 0.47 to 1).

There is strong evidence suggesting that there is a clinically important difference favouring St John's wort over antidepressants on:

- reducing the likelihood of patients leaving treatment early due to side effects (K = 10; N = 1629; RR = 0.39; 95% CI, 0.26 to 0.6)
- reducing the likelihood of patients reporting adverse effects (K = 8; N = 1358; RR = 0.65; 95% CI, 0.57 to 0.75).

7.10.4 Clinical summary

St John's wort is more effective than placebo on achieving response in both moderate and severe depression, and on reducing symptoms of depression in moderate depression.

There appears to be no difference between St John's wort and other antidepressants, other than in moderate depression where it is better at achieving response and in severe depression where it is less effective than low-dose antidepressants in achieving response.

However, St John's wort appears as acceptable as placebo and more acceptable than antidepressants, particularly TCAs, with fewer people leaving treatment early due to side effects and reporting adverse events.

7.10.5 Recommendations

74. Although there is evidence that St John's wort may be of benefit in less severe depression, practitioners should:

- not prescribe or advise its use by people with depression because of uncertainty about appropriate doses, persistence of effect, variation in the nature of preparations and potential serious interactions with other drugs (including oral contraceptives, anticoagulants and anticonvulsants)
- advise people with depression of the different potencies of the preparations available and of the potential serious interactions of St John's wort with other drugs [2004].

7.11 Seasonal affective disorder

7.11.1 Databases searched and the inclusion/exclusion criteria

Information about the databases searched for published trials and the inclusion/exclusion criteria used are presented in Table 81 . Details of the search strings used are in Appendix H.

Table 81: Databases searched and inclusion/exclusion criteria for clinical effectiveness of psychological treatments

Electronic databases	MEDLINE, EMBASE, PsycINFO, CINAHL
Date searched	Database inception to January 2008
Update searches	July 2008; January 2009
Study design	RCT
Population	People with a diagnosis of depression with a seasonal pattern according to DSM, ICD or similar criteria, or seasonal affective disorder according to Rosenthal's (1984) criteria or subsyndromal major depression with a seasonal pattern as indicated by score on seasonal depression scale
Treatments	Light therapy, dawn simulation, antidepressants, psychological therapies, other physical treatments

7.11.2 Light therapy for depression with a seasonal pattern

Depression with a seasonal pattern was not included in the scope of the previous guideline. Light therapy, which has been developed as a treatment specifically for major depression with a seasonal pattern, was therefore not reviewed, but has been included here as an additional review for the guideline update. For this review both published and unpublished RCTs investigating light therapy in patients diagnosed with major or subsyndromal major depression with a seasonal pattern were sought. There are a range of methods for administering light therapy; this review included a range of light treatments such as a light box, light room or visor and dawn simulation. Trials comparing a light treatment with a control condition, another light treatment or light administered at different times of day were included in this review.

A special adviser was consulted regarding a number of issues for this review (see Appendix 3). He advised the GDG that 5,000 lux hours[n] per day is a reasonable minimum dose for light box treatment, but that a minimum effective dose of light administered by a light visor has not yet been established. For the control light condition a placebo light of not more than 300 lux is appropriate. He suggested that a mini- mum trial duration of a week would be reasonable for evaluating the efficacy of light treatment. His advice was also sought regarding dawn simulation; he suggested that it would be informative to include this type of light treatment in the review and that a simulation of around an hour and a half peaking at 250 lux is an appropriate minimum, with a control condition of a light of less than 2 lux.

7.11.2.1 Studies considered[o]

In total, 61 trials were found from searches of electronic databases. Of these, 19 were included and 42 were excluded. The most common reasons for exclusion were that papers were not RCTs or participants did not have a diagnosis of depression or subsyndromal depressive symptoms with a seasonal pattern. In addition, studies that used a cross-over

n Lux is a standard measure of illuminance; 1 lux is equal to 1 lumen per square metre [lumen is the unit of luminous flux].

o Study IDs in capital letters refer to studies found and included in this guideline update.

design (where participants serve as their own controls by receiving both treatments) were not used unless pre-crossover data were available.

The studies that were found by the search and included in this review varied considerably in methodology. The intensity and duration of light, time of day, mode of administration of light, and the comparison conditions were different across studies. A range of outcomes were reported by the included studies, including the HRSD (termed 'typical' depression rating scale to distinguish it from scales measuring depression with seasonal pattern symptoms), and scales adapted for measuring symptoms in depression with a seasonal pattern. These included the Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH) for major depression with a seasonal pattern (Williams *et al.*, 1988), which combines the HRSD with an additional eight items relevant to depression with a seasonal pattern. Some studies report the eight additional items separately. Both typical and atypical symptoms were measured using clinician- and self-rated scales. All data were extracted and can be seen in the full evidence profiles and forest plots (Appendix J11 and Appendix L, respectively). Only data for the SIGH for major depression with a seasonal pattern (clinician- and self-rated) are presented here.

Data were available to compare light therapy with a range of control conditions including waitlist, attentional controls and active treatment controls. In addition administration of light in the morning versus evening was compared and dawn simulation was compared with attentional control and with bright light. One study included a combination treatment of light and CBT and one trial reported on light therapy for relapse prevention.

Summary study characteristics of the included studies are presented in Table 82 and Table 83 with full details in Appendix J11, which also includes details of excluded studies.

Table 82: Summary study characteristics of light therapy studies versus control and morning light versus afternoon/evening light

	Light versus waitlist control	Light versus attentional control	Light versus active treatment control	Morning versus afternoon/evening light
No. trials (total participants)	2 RCTs (82)	8 RCTs (401)	4 RCTs (243)	4 RCTs (144)
Study IDs	RASTAD2008 ROHAN2007	DESAN2007 EASTMAN1998 JOFFE1993 LEVITT1996 ROSENTHAL1993 STRONG2008 TERMAN1998† WILEMAN2001	LAM2006F MARTINEZ1994 ROHAN2004 ROHAN2007	AVERY2001A EASTMAN1998 LAFER1994‡ TERMAN1998†
N/% female	(1) 51/80 (2) 31/84	(1) 26/77 (2) 81/88 (3) 67/87 (4) 44/72 (5) 55/84 (6) 30/78 (7) 39/80 (8) 59/88	(1) 96/67 (2) 20/65 (3) 26/92 (4) 61/94	(1) 31/90 (2) 81/85 (3) 32/65 (4) 39/80
Mean age	(1) 46 (2) 45	(1) 46 (2) 37 (3) 40	(1) 43 (2) 46 (3) 51	(1) 40 (2) 37 (3) 35

	Light versus waitlist control	Light versus attentional control	Light versus active treatment control	Morning versus afternoon/evening light
		(4) 35 (5) 42 (6) 44 (7) 39 (8) 41	(4) 45	(4) 39
Diagnosis	(1)–(2) MDD with seasonal pattern (DSM–IV)	(1) MDD with seasonal pattern (DSM–IV) (2) Major depression with a seasonal pattern (Rosenthal) (3) MDD or bipolar with seasonal pattern (DSM–III-R) or major depression with a seasonal pattern (Rosenthal) (4) MDD with seasonal pattern (DSM–III-R) (5) Major depression with a seasonal pattern (Rosenthal) (6) MDD with seasonal pattern (DSM–IV) (7) Mood disorder with major depression with a seasonal pattern (DSM–III-R) (8) MDD with seasonal pattern (DSM–IV)	(1) MDD or bipolar with seasonal pattern (DSM–IV) (2) MDD with seasonal pattern (DSM–III-R) (3)–(4) MDD with seasonal pattern (DSM–IV)	(1) Subsyndromal major depression with a seasonal pattern (2) Major depression with a seasonal pattern (Rosenthal) (3) Major depressive episode with a seasonal pattern (DSM–III-R) (4) Mood disorder with major depression with a seasonal pattern (DSM–III-R)
Light therapy	(1) Fluorescent light room (2) Fluorescent light box	(1) LED Litebook device (2) Fluorescent light box (3) Light visor (4a) Fluorescent light box (4b) LED visor (5) Light visor (6) Narrow-band blue light panel (7)–(8) Light box	(1) Fluorescent light box + placebo pill (2) Light box + hypericum (3) Light box (4) Fluorescent light box	(1) Light box used between 7 am–12 pm (2) Fluorescent light box used as soon as possible after waking (3) Bright light for 2 hours (4) Light box 10 minutes after waking
Lux hours/day	(1) Varies 1650–8600	(1) 675 (2) 9000 (3) Mean 1762	(1) 5000 (2) 3000 (3) 15000	(1) 5000 (2) 9000 (3) 2,500

	Light versus waitlist control	Light versus attentional control	Light versus active treatment control	Morning versus afternoon/evening light
	(2) 15000 in 1st week, varies after week 1	(4a) Mean 3800 (4b) Mean 323 (5) 3000 or 6000 (6) 470 nm 176 lux X 45 minutes (7) 10000 (8) 5000 in 1st week, 7500 in 2nd week, 10000 in last 2 weeks	(4) 15000 in 1st week, varies after week 1	(4) 10000
Comparator(s)	(1)–(2) Waitlist	(1)–(2) Deactivated negative ion generator (3) Dim 67 lux light visor (4a) Light box producing no light (4b) Visor producing no light (5) Dim 400 lux light visor (6) Red light (7) Low-density negative ions (8) Dim 500 lux red light box	(1) Dim 100 lux light + 20 mg/day fluoxetine (2) Dim light + hypericum (3) Group CBT/light + group CBT (4) Group CBT	(1) Light box used between 12–5 p.m. (2) Fluorescent light box used within 1 hour of bedtime (3) Bright light for 2 hours (4) Light box 2–3 hours before bedtime
Length of treatment (days)	(1) 21 (2) 42	(1)–(2) 28 (3)–(4) 14 (5) 7 (6) 21 (7) 14 (8) 28	(1) 56 (2) 28 (3)–(4) 42	(1) 14 (2) 28 (3) 7 (4) 14

**3-armed trial, †5-armed trial and ‡3-armed trial but 1 arm not used (bright light alternating morning and evening).*

Table 83: Summary study characteristics of dawn simulation and relapse prevention studies

	Dawn simulation versus attentional control	Light versus dawn simulation	Relapse prevention
No. trials (total participants)	3 RCTs (139)	2 RCTs (112)	1 RCT (46)
Study IDs	AVERY1993 AVERY2001 TERMAN2006	AVERY2001 TERMAN2006	(1) MEESTERS 1999
N/% female	(1) 27/70	(1) 64/88	(1) 46/71
	(2) 62/87	(2) 48	
	(3) 50/79		
Mean age	(1) 35	(1) 41	(1) 40
	(2) 41	(2) 40	

	Dawn simulation versus attentional control	Light versus dawn simulation	Relapse prevention
	(3) 40		
Diagnosis	Major depression with a seasonal pattern (Rosenthal) MDD or bipolar with seasonal pattern (DSM–IV) MDD with seasonal pattern (DSM–III-R)	MDD or bipolar with seasonal pattern (DSM–IV) MDD with seasonal pattern (DSM–III-R)	(1) MDD with seasonal pattern (DSM–IV)
Light therapy	Gradual dawn simulation over 2 hours Gradual dawn simulation over 1.5 hours (3) Gradual dawn simulation over 3.5 hours	(1)–(2) Light box	(1) Light visor
Lux hours/day	(1)–(3) 250 lux peak intensity	(1) 5000 (2) 10000	(1) 1250
Comparator	(1) Rapid dim 0.2 lux dawn Dim 0.5 lux red dawn Pulse dawn 250 lux 30 minutes	Gradual dawn simulation over 1.5 hours peaking at 250 lux Gradual dawn simulation over 3.5 hours	(1a) No treatment (1b) Dim 0.18 lux infrared light
Length of	(1) 7	(1) 42	(1) 182
treatment (days)	(2) 42	(2) 21	
	(3) 21		

7.11.31 Clinical evidence

7.11.3.12 Bright light versus waitlist or attentional control

Compared with waitlist control, bright light (either light room or light box) shows a strong effect on symptoms in depression with a seasonal pattern although there are few studies. Compared with attentional controls, such as deactivated negative ion generator, dim red light, and sham light boxes, bright light (either via light box or light visor) shows a small effect on symptoms in depression with a seasonal pattern that was not clinically important. Evidence from the important outcomes and overall quality of evidence are presented in Table 84. The full evidence profiles and associated forest plots can be found in Appendix J11 and Appendix L, respectively.

7.11.3.21 Bright light versus active treatment control

There were data to compare light therapy with group CBT, light therapy plus CBT, and dim light plus fluoxetine. There was also a study comparing light therapy plus St John's wort with dim light plus St John's wort.

Compared with group CBT (tailored to depression with a seasonal pattern) bright light therapy was no better in terms of reducing depressive symptoms in depression with a seasonal pattern, although the effect size is not statistically significant and was graded low quality. However, more participants achieved remission with bright light therapy than with

group CBT (52% compared with 37.5%), although the result is not clinically important. Similarly, light therapy appeared to be more acceptable than group CBT with fewer people leaving treatment early (8% compared with 16.7%) although the effect size is not statistically significant. Treatment lasted for 6 weeks.

Combination treatment (bright light plus CBT) was more effective than light therapy alone on both the SIGH for major depression with a seasonal pattern and the BDI, although the effect sizes were not statistically significant. Roughly equal numbers of participants left treatment early.

There appeared to be little difference between bright light therapy and fluoxetine (20 mg) on efficacy outcomes (both treatments given with a sham treatment mimicking the other). Treatment lasted for 8 weeks.

There was no evidence for the efficacy of light therapy combined with St John's wort compared with a sham light condition plus St John's wort. There was only a single small 4-week study (n = 20).

Evidence from the important outcomes and overall quality of evidence are presented in Table 85. The full evidence profiles and associated forest plots can be found in Appendix J11 and Appendix L, respectively.

7.11.48 Morning light versus afternoon/evening light

Three studies compared light therapy administered in the morning compared with light therapy in the afternoon or evening, one of which was in participants with subsyndromal major depression with a seasonal pattern. There were no significant differences in outcome measures for those given light therapy in the morning compared with those given light therapy in the afternoon or evening. Evidence from the important outcomes and overall quality of evidence are presented in Table 86. The full evidence profiles and associated forest plots can be found in Appendix J11 and Appendix L, respectively.

Table 84: Summary evidence profile for bright light versus waitlist or attentional controls

	Bright light versus waitlist control	Bright light versus attentional control
Leaving treatment early	RR 0.95 (0.21 to 4.32) (7.1 versus 7.5%)	RR 0.88 (0.50 to 1.54) (13.4 versus 14.5%)
Quality	Low	Low
Number of studies; participants	K = 2; n = 82	K = 6; n = 266
Forest plot number	Pharm SAD 01.01	Pharm SAD 02.01
Reported side effects	Not reported	RR 0.98 (0.73 to 1.32) (55.6 versus 58.3%)
Quality	–	Low
Number of studies; participants	–	K = 2; n = 81
Forest plot number	–	Pharm SAD 02.03
Clinician-rated endpoint (SIGH-SAD)	WMD -10.4 (-15.99 to -4.81)	WMD -3.07 (-6.71 to 0.58)
Quality	Moderate	Low
Number of studies; participants	K = 1; n = 31	K = 8; n = 300
Forest plot number	Pharm SAD 01.04	Pharm SAD 02.04
Self-rated endpoint (SIGH-SAD-SR)	WMD -12.8 (-18.52 to -7.08)	Not reported

	Bright light versus waitlist control	Bright light versus attentional control
Quality	Moderate	–
Number of studies; participants	K = 1; n = 44	–
Forest plot number	Pharm SAD 01.03	–
Non-remission (based on SIGH-SAD-SR)	RR 0.53 (0.38 to 0.74) (47.6 versus 90%)	RR 0.89 (0.66 to 1.2) (56.3 versus 61.3%)
Quality	High	Low
Number of studies; participants	K = 2; n = 82	K = 6; n = 336
Forest plot number	Pharm SAD 01.10	Pharm SAD 02.08
Non-response (based on SIGH-SAD	RR 0.50 (0.34 to 0.73) (50 versus 100%)	RR 0.86 (0.64 to 1.15) (45.4 versus 53.8%)
Quality	Moderate	Low
Number of studies; participants	K = 1; n = 51	K = 7; n = 354
Forest plot number	Pharm SAD 01.11	Pharm SAD 02.09

1 **Table 85: Summary evidence profile for bright light versus active treatment control**

	Light box versus group CBT	Light box versus light box + group CBT	Light box + placebo pill versus dim light box + fluoxetine	Light box + St John's wort versus dim light + St John's wort
Leaving treatment early	RR 0.53 (0.12 to 2.31) (8 versus 16.7%)	RR 0.92 (0.17 to 4.91) (8 versus 8.7%)	RR 1.14 (0.45 to 2.90) (16.7 versus 14.6%)	Not reported
Quality	Moderate	Moderate	Moderate	–
Number of studies; participants	K = 2; n = 49	K = 2; n = 48	K = 1; n = 96	–
Forest plot number	Pharm SAD 03.01	Pharm SAD 04.01	Pharm SAD 03.01	–
Reported side effects	Not reported	Not reported	RR 1.03 (0.82 to 1.29) (77.1 versus 75%)	Not reported
Quality	–	–	Moderate	–
Number of studies; participants	–	–	K = 1; n = 96	–
Forest plot number	–	–	Pharm SAD 03.04	–
Clinician-rated mean endpoint	WMD -0.2 (-6.5 to 6.1) (SIGH-SAD)	WMD 4.2 (-0.52 to 8.92) (SIGH-SAD)	WMD -0.00 (-3.88 to 3.88) (SIGH-SAD)	SMD -0.32 (-1.2 to 0.57) (HRSD)
Quality	Low	Moderate	High	Low
Number of studies; participants	K = 1; n = 31	K = 1; n = 31	K = 1; n = 96	K = 1; n = 20
Forest plot number	Pharm SAD 03.05	Pharm SAD 04.03	Pharm SAD 03.05	Pharm SAD 03.06

	Light box versus group CBT	Light box versus light box + group CBT	Light box + placebo pill versus dim light box + fluoxetine	Light box + St John's wort versus dim light + St John's wort
Self-rated mean endpoint	WMD -0.7 (-7.16 to 5.76) (BDI)	SMD 2.3 (-2.47 to 7.07) (BDI)	WMD -1.6 (-5.68 to 2.48) (BDI)	Not reported
Quality	Low	Low	Low	–
Number of studies; participants	K = 1; n = 31	K = 1; n = 31	K = 1; n = 96	–
Forest plot number	Pharm SAD 03.08	Pharm SAD 04.06	Pharm SAD 03.08	–
Non-remission (based on SIGH-SAD-SR)	RR 0.77 (0.46 to 1.28) (48 versus 62.5%)	RR 2.22 (0.92 to 5.32) (48 versus 21.7%)	RR 1.09 (0.57 to 1.76) (50 versus 45.8%)	Not reported
Quality	High	High	Low	–
Number of studies; participants	K = 2; n = 49	K = 2; n = 48	K = 1; n = 96	–
Forest plot number	Pharm SAD 03.09	Pharm SAD 04.07	Pharm SAD 03.09	–
Non-response (based on SIGH-SAD-SR)	Not reported	Not reported	RR 1 (0.57 to 1.76) (33.3 versus 33.3%)	Not reported
Quality	–	–	Low	–
Number of studies; participants	–	–	K = 1; n = 96	–
Forest plot	–	–	03.10	–

1 **Table 86: Summary evidence profile for morning light versus evening light**

	Overall results	Subsyndromal major depression with a seasonal pattern only
Leaving treatment early	RR 0.98 (0.41 to 2.35) (12.1 versus 12.5%)	Not reported
Quality	Moderate	–
Number of studies; participants	K = 3; n = 130	–
Forest plot number	Pharm SAD 05.01	–
Reported side effects	RR 0.47 (0.05 to 4.65) (6.3 versus 13.3%)	RR 0.47 (0.05 to 4.65) (6.3 versus 13.3%)
Quality	Low	Low
Number of studies; participants	K = 1; n = 31	K = 1; n = 31
Forest plot number	Pharm SAD 05.03	Pharm SAD 05.03
Clinician-rated mean endpoint	WMD -1.38 (-5.49 to 2.73) (SIGH-SAD)	WMD 0.6 (-3.89 to 5.09) (SIGH-SAD)
Quality	Low	Low

	Overall results	Subsyndromal major depression with a seasonal pattern only
Number of studies; participants	K = 2; n = 68	K = 1; n = 30
Forest plot number	Pharm SAD 05.04	Pharm SAD 05.04
Self-rated mean endpoint	WMD -0.9 (-4.66 to 2.86) (BDI)	Not reported
Quality	Low	–
Number of studies; participants	K = 1; n = 65	–
Forest plot number	Pharm SAD 05.07	–
Non-remission (based on SIGH-SAD-SR)	RR 1.0 (0.69 to 1.45) (54 versus 54.2%)	Not reported
Quality	Low	–
Number of studies; participants	K = 2; n = 98	–
Forest plot number	Pharm SAD 05.08	–
Non-response (based on SIGH-SAD-SR)	RR 1.0 (0.51 to 1.98) (44 versus 42.9%)	RR 0.52 (0.23 to 1.20) (31.3 versus 60%)
Quality	Low	Moderate
Number of studies; participants	K = 3; n = 129	K = 1; n = 31
Forest plot number	Pharm SAD 05.09	Pharm SAD 05.09

7.11.4.11 Dawn simulation versus attentional control or light therapy

Three studies compared dawn simulation with an attentional control. There was some evidence that dawn simulation improved symptoms of depression but it was not clinically important and was not supported by other outcomes including the major depression with a seasonal pattern subscale. Similarly, there was no evidence of superiority of dawn simulation over regular light therapy. Evidence from the important outcomes and overall quality of evidence are presented in Table 87. The full evidence profiles and associated forest plots can be found in Appendix J11 and Appendix L, respectively.

Table 87: Summary evidence profile for dawn simulation studies

	Dawn simulation versus attentional control	Light therapy versus dawn simulation
Leaving treatment early	RR 0.27 (0.08 to 0.92) (2.9 versus 14.1%)	RR 3.72 (0.62 to 22.22) (8.9 versus 1.8%)
Quality	Low	Moderate
Number of studies; participants	K = 3; n = 141	K = 2; n = 112
Forest plot number	Pharm SAD 06.01	Pharm SAD 07.01
Reported side effects	RR 5.57 (0.77 to 40.26) (42.9 versus 7.7%)	Not reported
Quality	Low	–
Number of studies; participants	K = 1; n = 27	–
Forest plot number	Pharm SAD 06.04	–
Clinician-rated mean endpoint	SMD -0.53 (-1.62 to 0.15) (HRSD) WMD -2.20 (-7.52 to 3.11) (SAD subscale)	WMD -0.9 (-4 to 2.2) (HRSD) WMD -1.8 (-6.98 to 3.38) (SAD subscale)

	Dawn simulation versus attentional control	Light therapy versus dawn simulation
Quality	Moderate (HRSD) Very low (SAD subscale)	Very low (HRSD) Low (SAD subscale)
Number of studies; participants	K = 2; n = 73	K = 1; n = 45
Forest plot number	Pharm SAD 06.05/06	Pharm SAD 07.06/07
Self-rated mean endpoint	Not reported	Not reported
Quality	–	–
Number of studies; participants	–	–
Forest plot number	–	–
Non-remission (based on SIGH-SAD)	RR 0.9 (0.46 to 1.78) (44.6 versus 50%)	RR 1.19 (0.70 to 2.00) (53.6 versus 44.6%)
Quality	Low	Very low
Number of studies; participants	K = 2; n = 114	K = 2; n = 112
Forest plot number	Pharm SAD 06.07	Pharm SAD 07.04
Non-response (based on SIGH-SAD)	RR 0.71 (0.34 to 1.48) (25 versus 38%)	RR 1.45 (0.82 to 2.58) (35.7 versus 25%)
Quality	Moderate	Moderate
Number of studies; participants	K = 2; n = 114	K = 2; n = 112
Forest plot number	Pharm SAD 06.08	Pharm SAD 07.05

Prevention of future episodes using light therapy

One study compared bight light therapy with a control treatment and with no treatment as relapse prevention in people who had a history of depression with a seasonal pattern but had not yet developed symptoms. This showed that those receiving light therapy were less likely to develop symptoms of depression compared with those receiving no treatment. However, those using the infrared light visor were less likely to develop symptoms of depression than those using the bright white light visor. Neither finding was clinically important. Evidence from the important outcomes and overall quality of evidence are presented in Table 88. The full evidence profiles and associated forest plots can be found in Appendix J11 and Appendix L, respectively.

Table 88: Summary evidence profile for relapse prevention using bright light

	Bright white light visor versus no treatment control	Bright white light visor versus infrared light visor
Leaving treatment early	RR 2.22 (0.29 to 17.27) (22.2 versus 10%)	RR 1.33 (0.35 to 5.13) (22.2 versus 16.7%)
Quality	Low	Low
Number of studies; participants	K = 1; n = 28	K = 1; n = 36
Forest plot number	Pharm SAD 08.01	Pharm SAD 08.01
Relapse (BDI >13 for 2 consecutive weeks)	RR 0.63 (0.36 to 1.09) (50 versus 80%)	RR 2.25 (0.84 to 5.99) (50 versus 22.2%)
Quality	Moderate	Moderate
Number of studies; participants	K = 1; n = 28	K = 1; n = 36
Forest plot number	Pharm SAD 08.02	Pharm SAD 08.02

7.11.4.2 Clinical summary

Although there are a large number of studies that address the efficacy of light treatment in people with depression that follows a seasonal pattern, these studies are difficult to interpret due to methodological differences. The doses and colours of light, methods of delivery, comparator treatments, and clinical populations included in studies are diverse. While bright light is clearly more effective than waitlist control, it is unclear if this is more than a placebo effect (see discussion on the placebo effect in Chapter 2, Section 2.4.3). Studies that compare bright light with other treatments that are not known to be effective give equivocal results. There are too few data relating to active controls to determine non-inferiority, and few systematic data relating to side effects. In clinical practice, where bright light is used, a minimum daily dose of 5,000 lux administered in the morning during the winter months is the most common treatment strategy. The most common side effect seen is mild agitation.

7.11.5 Other therapies for depression with a seasonal pattern

7.11.5.1 Studies considered[p]

In total, 14 trials of interventions other than bright light were found, mostly of anti-depressants, of which five met inclusion criteria for a review of acute-phase treatment, one for a review of continuation treatment in people who had responded to open-label treatment, and three (published in the same paper) for a review of prevention in people with a history of depression with a seasonal pattern. Summary study characteristics of the included studies are presented in Table 89, with full details in Appendix J11, which also includes details of excluded studies.

Table 89: Summary study characteristics for interventions other than bright light for major depression with a seasonal pattern

	Acute phase treatments	Continuation treatment	Prevention treatment
No. trials (total participants)	5 RCTs (346)	1 RCTs (23)	3 RCTs (1061)
Study IDs	(1) LAM1995 (2) LINGJAERDE1993 (3) MOSCOVITCH2004 (4) PARTONEN1996 (5) TERMAN1995	(1) SCHLAGER1994*	(1) MODELL2005 study 1 (2) MODELL2005 study 2 (3) MODELL2005 study 3
N/% female	(1) 68/66 (2) 34/74 (3) 187/78 (4) 32/66 (5) 25/88	(1) 23 (not available)	(1) 277/72 (2) 311/67 (3) 473/68
Mean age	(1) 36 (2) 43 (3) 40 (4) 44 (5) 38	(1) Not given	(1) 42 (2) 42 (3) 41
Diagnosis	(1) Recurrent major depressive episodes with seasonal pattern	(1) Responders to initial treatment for recurrent major	(1)–(3) History of MDD with seasonal pattern (DSM-IV)

p Study IDs in title case refer to studies included in the previous guideline and study IDs in capital letters refer to studies found and included in this guideline update. References for studies from the previous guideline are in Appendix U.

	Acute phase treatments	Continuation treatment	Prevention treatment
	(2) Mood disorder with seasonal pattern (3) 79% major depression with seasonal pattern; 13% depression NOS with seasonal pattern; 7% bipolar disorder with seasonal pattern; 2% bipolar disorder NOS with seasonal pattern (4) 100% MDD; 18% mood disorder with seasonal pattern (5) Major depression with a seasonal pattern, MDD with seasonal pattern, or bipolar disorder NOS with seasonal pattern - % not clear	depressive episodes with seasonal pattern	
Treatment	Fluoxetine 20 mg Moclobemide 400 mg Sertraline 50–200 mg Moclobemide 300–450 mg High density negative ions	(1) Propanolol 33 mg	(1) Buspirone 150–300 mg (2)–(3) Bupropion XL 150–300 mg
Comparator	(1)–(3) Placebo Fluoxetine 20–40 mg Low density negative ions	(1) Placebo	(1)–(3) Placebo
Length of treatment (days)	5 weeks 3 weeks 8 weeks 6 weeks 3 weeks	(1) 2 weeks	(1) 6 months (2)–(3) Unclear

**Continuation trial.*

7.11.5.22 Clinical evidence

Acute-phase treatments

The data for acute-phase treatment comparing antidepressants with placebo were largely inconclusive, although on one outcome (response) there appeared to be little difference. Acceptability and tolerability data were inconclusive. There was no evidence to suggest a difference between moclobemide and fluoxetine, which was the only head-to-head evidence available. There was some evidence to suggest that high ion density was more effective than low ion density, although there was only one study. Evidence from the important outcomes and overall quality of evidence are presented in Table 90. The full evidence profiles and associated forest plots can be found in Appendix J11 and L, respectively.

Table 90: Summary evidence profile for acute-phase treatments (not light therapy) for major depression with a seasonal pattern

	Antidepressants versus placebo	Antidepressants versus antidepressants	High ion density versus low ion density
Non-response (based on SIGH-SAD)	RR 0.82 (0.63 to 1.05) (44.2 versus 54%)	Not reported	RR 0.49 (0.24 to 1) (41.7 versus 84.6%)
Quality	High	–	Moderate
Number of studies; participants	K = 2; n = 255	–	K = 1; n = 25
Forest plot number	Pharm SAD 09.01	–	Pharm SAD 12.01
Clinician-rated mean endpoint SIGH-SAD	SMD -0.11 (-0.65 to 0.42)	Moclobemide versus fluoxetine: WMD -1.6 (-7.01 to 3.81)	Not reported
Quality	Low	Low	–
Number of studies; participants	K = 2; n = 99	K = 1; n = 29	–
Forest plot number	Pharm SAD 09.02	Pharm SAD 11.01	–
Self-rated mean endpoint BDI	WMD -1.7 (-6.53 to 3.13)	Not reported	Not reported
Quality	Low	–	–
Number of studies; participants	K = 1; n = 68	–	–
Forest plot number	Pharm SAD 09.02	–	–
Leaving treatment early	RR 0.7 (0.16 to 3.05) (18.3 versus 20.5%)	Not reported	Not reported
Quality	Very low	–	–
Number of studies; participants	K = 2; n = 221	–	–
Forest plot number	Pharm SAD 10.01	–	
Leaving treatment early due to side effects	RR 1.48 (0.63 to 3.47) (8.3 versus 5.6%)	Not reported	Not reported
Quality	Low	–	–
Number of studies; participants	K = 3; n = 289	–	–
Forest plot number	Pharm SAD 10.02	–	–

Continuation treatment and prevention of future episodes

One small study compared the [3-blocker, propanolol, with placebo for people who had responded to previous open treatment. This showed that symptoms of depression in those continuing treatment remained lower compared with those switched to placebo. Another three trials compared bupropion with placebo to prevent episodes in people with a history of depression. Treatment started before the onset of winter and continued until early spring. There was a clinically important reduction in the number of recurrences among those taking bupropion compared with the rate in those taking placebo. Evidence from the important outcomes and overall quality of evidence are presented in Table 91. The full evidence profiles and associated forest plots can be found in Appendix J11 and Appendix L, respectively.

Table 91: Summary evidence profile of continuation treatment and prevention of future episodes for people with major depression with a seasonal pattern

	Continuation treatment: propanolol versus placebo	Prevention: bupropion versus placebo
Efficacy outcome	HAMD-21: WMD -7 (-11.24 to -2.76)	Recurrence: RR 0.58 (0.46 to 0.72) (17% versus 29.5%)
Quality	Moderate	High
Number of studies; participants	K = 1; n = 23	K = 3; n = 1061
Forest plot number	Pharm SAD 13.01	Pharm SAD 14.01
Leaving treatment early	RR 2.57 (0.12 to 57.44) (7.7 versus 0%)	Not reported
Quality	Low	–
Number of studies; participants	K = 1; n = 24	–
Forest plot number	Pharm SAD 13.02	–

7.11.5.3 Clinical summary

There was a lack of evidence for the effectiveness of antidepressants in the treatment of major depression with a seasonal pattern once symptoms have begun but evidence for a prophylactic effect of starting treatment before symptoms start and continuing until early spring.

7.11.6 From evidence to recommendations

The evidence for light therapy for major depression with a seasonal pattern is poorly developed, with many trials comparing different elements of treatment, including time of day, level of light and length of treatment. There is little evidence for the efficacy of bright light in the treatment of major depression with a seasonal pattern compared with placebo treatment.

The evidence for other treatments is sparse. Evidence is lacking that antidepressants are effective once symptoms have begun, but they may be worthwhile as prophylactics. For depression with a seasonal pattern practitioners should follow the guidance for depression elsewhere in this guideline.

7.11.7 Recommendations

75. Advise people with winter depression that follows a seasonal pattern and who wish to try light therapy in preference to antidepressant or psychological treatment that the evidence for the efficacy of light therapy is uncertain. [2009]

8 Further-line treatment of depression

8.1 Introduction

8.1.1 Failure of first-line treatment

Adequate first-line treatments for depression are associated with non-remission in roughly two-thirds of cases (Rush et al. 2006). The question of what to do following treatment failure is therefore a common clinical dilemma for patients and professionals. Common further-line treatment strategies include switching to a different medication or to psychotherapy. Choice of second-line strategy is often informed by preference and availability, although patient characteristics including previous history of treatment response, type of depressive syndrome and co-morbidities can be helpfully used to guide the next step.

For the substantial proportion of patients who remain in depression following second-line treatment failure, the terms 'treatment resistance' or 'treatment resistant depression' (TRD) are often used.

8.1.2 Treatment resistance

'Treatment resistance' is generally considered as a failure to respond to 2 adequate courses of antidepressants within a specified episode of depression (Burrows et al. 1994, Souery et al. 1999, Souery et al. 2006). Over the last 20 years there have been a number of attempts to operationalise this concept further, with controversy over the best way to measure the degree of resistance to treatment. An early attempt at 'staging' treatment resistance incorporated both the number of treatments attempted and a hierarchy of treatments; including for example the failure of treatment with tricyclic antidepressants (stage III resistance) at a lower level than failure with mono-amine oxidase inhibitors (stage IV resistance) (Thase and Rush 1997). Whilst evidence supports the first part of this model (absolute numbers of treatment failures), since rates of remission drop sharply after the first 2 treatment attempts (from around 30% to less than 15%) (Rush et al. 2006), there is much less robust evidence for the superiority of one agent over another in treatment resistance (for example, tricyclics versus venlafaxine) and therefore the hierarchical aspect has been challenged (Fava 2003).

More recent models (such as the Massachusetts General Hospital [Fava 2003] and the Maudsley Staging Method [Fekadu et al. 2009]) have sought to avoid the idea of a hierarchy of antidepressants; to specify the dose and duration of antidepressant treatment that can be considered adequate; and to account for the failure of combination and augmentation strategies (in addition to trials of single antidepressant agents). A systematic review of all of these approaches identified that the Maudsley Staging Method had the best predictive utility in assessing resistance (Ruhe et al. 2012). However, all of these staging methods remain limited through their focus on assessing resistance to biological treatments within the current episode. Recent clinical trials (Keller et al. 2000, Thase et al. 2007, Kocsis et al. 2009, Wiles et al. 2013) and functional neuroimaging studies (McGrath et al. 2013) have suggested that some types of psychotherapy may have an important place in overcoming treatment resistance. Further clarifying this role, particularly at later stages of treatment failure, may help in developing fuller models of treatment resistance and likelihood of future remission.

Alongside efforts to more clearly delineate treatment resistance there has been greater acknowledgement of so-called 'pseudo-resistance', where lack of response relates to misdiagnosis (for example, of bipolar depression) or undertreatment (for example, through inadequate dosage or length of treatment [Keller et al. 1995]), rather than true treatment resistance. Understanding this problem of 'pseudo-resistance' (and avoiding incorrectly

Update 2017

labelling an individual as genuinely treatment resistant) should remain a significant concern in day-to-day clinical practice in order to improve treatment outcomes.

Genuine treatment resistance has been linked to a number of demographic and illness characteristics, including: living alone; lower income; unemployment; male gender; lower education; higher complexity through associated physical or psychiatric disorder; and a longer, more severe current episode (Trivedi et al. 2006). Several approaches to overcoming resistant depression have been evaluated, including pharmacology, neurostimulation and psychotherapy. Pharmacological next-step options include: switching within a class of antidepressants (for example, different SSRIs); switching between different classes of antidepressants (for example, from an SSRI to a SNRI); combining different antidepressants together (for example, SSRI plus mirtazapine); or augmenting an antidepressant with an agent that is not antidepressant in its own right (for example, lithium). Given the lack of convincing superiority of one agent over another at group level, part of the therapeutic advantage of switching between antidepressants may come through 'pharmacogenomics', indicating the genetic factors that may make people differentially liable to the beneficial or adverse effects of particular pharmacological agents (Perlis 2014, Coplan et al. 2014).

Evidence indicates that people continue to achieve remission when further treatment steps are used but that even with this approach around one third of people will remain treatment resistant at one year (Rush et al. 2006). After a period of treatment resistance there is some evidence that remission is less stable, associated with higher subsequent relapse and shorter average time to relapse (Rush et al. 2006); indicating over the longer term that those people who find it difficult to get well may also then find it more difficult to stay well.

8.23 Review questions

Update 2017

- For adults with depression following no or limited response to previous treatment (of the current episode), or those not tolerating previous treatment (of the current episode), what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination?
- For adults with treatment-resistant depression, what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination?

The review protocol summary and the eligibility criteria used for this section of the guideline, can be found in Table 92. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 92: Clinical review protocol summary for the review of further-line treatment

Component	Description
Review questions	For adults with depression following no or limited response to previous treatment (of the current episode), or those not tolerating previous treatment (of the current episode), what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination? (RQ2.4) For adults with treatment-resistant depression, what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination? (RQ2.5)
Population	Adults in a depressive episode whose depression has not responded or there has been limited response to previous treatment(s) (for the current episode) according to DSM, ICD or similar criteria, or (residual) depressive symptoms as indicated by depression scale score, or who have not tolerated previous treatment (for the current episode), and who have been randomised to the further line

Component	Description
	interventions at the point at which they had no/adequate/limited response If some, but not all, of a study's participants are eligible for the review, and we are unable to obtain the appropriate disaggregated data, then we will include a study if at least 80% of its participants are eligible for this review
Intervention(s)	The following interventions will be included (alone, in combination or as augmentation strategies): **Psychological interventions**: • cognitive and cognitive behavioural therapies (including CBT, Mindfulness-based Cognitive Therapy [MBCT] and Cognitive Behavioural Analysis System of Psychotherapy [CBASP]) • counselling • interpersonal psychotherapy (IPT) • psychodynamic psychotherapy • self-help (with or without support) **Psychosocial interventions**: • befriending • peer support **Pharmacological interventions** • antidepressants ○ SSRIs - citalopram - escitalopram - fluvoxamine - fluoxetine - paroxetine - sertraline ○ TCAs - amineptine[1] - amitriptyline - clomipramine - desipramine[2] - imipramine - lofepramine - nortriptyline ○ TeCAs - mianserin ○ SNRIs - duloxetine - venlafaxine ○ other antidepressant drugs - bupropion[3] - mirtazapine • anticonvulsants ○ lamotrigine[3] • antipsychotics ○ amisulpride[3] ○ aripiprazole[3] ○ olanzapine[3] ○ quetiapine

Update 2017

Component	Description
	○ risperidone[3] ○ ziprasidone[2] • anxiolytics ○ buspirone • stimulants ○ methylphenidate[3] • other agents ○ lithium ○ omega-3 fatty acids ○ thyroid hormone[3] **Physical interventions** • ECT • exercise (including yoga) Interventions will be categorised into the following strategies: • dose escalation strategies • switching strategies (including switching to another antidepressant of the same class, switching to another antidepressant of a different class, and switching to a non-antidepressant treatment) • augmentation strategies (including augmenting the antidepressant with another antidepressant, augmenting the antidepressant with a non-antidepressant agent and augmenting the antidepressant with a psychological intervention)
Comparison	• Treatment as usual • Waitlist • Placebo • Any other active comparison In addition to placebo and head-to-head comparators, comparator treatment strategies include: • Continuing with the antidepressant at the same dose • Continuing with the antidepressant-only
Critical outcomes	**Efficacy** • Depression symptomology (mean endpoint score or change in depression score from baseline) • Remission (usually defined as a cut off on a depression scale) • Response (e.g. reduction of at least 50% from the baseline score on HAMD/MADRS) • **Acceptability/tolerability** • Discontinuation due to any reason (including adverse events) • Discontinuation due to adverse events • The following depression scales will be included: • MADRS • HAMD • QIDS • PHQ • CGI • CES-D • BDI • HADS-D (depression subscale) • HADS (full scale)

Update 2017

Component	Description
Study design	• RCTs • Cluster RCTs

Note: [1]Amineptine is not available to prescribe as a medicine (although it falls under Class C of the Misuse of Drugs Act 1971, and listed as Schedule 2 under the Controlled Drugs Regulations 2001). However, this drug is included in this review in order to assess the class effect of pharmacological interventions for depression
[2]Desipramine and ziprasidone are not available in the UK to prescribe. However, these drugs are included in this review in order to assess the class effect of pharmacological interventions for depression
[3]None of these drugs are licensed for use in depression. However, they are included in the review in order to assess harms and efficacy for off-label use and to assess the class effect of pharmacological interventions for depression

8.3 Clinical evidence

Two hundred and seven studies of further-line treatment for depression in adults were identified for full-text review. Of these 207 studies, 64 RCTs were included (Appelberg 2001; Barbee 2011; Bauer 2009; Bauer 2010/2013; Baumann 1996; Berman 2007; Berman 2009; Bose 2012; Browne 1990; Carpenter 2002; Chaput 2008; Chiesa 2015; Corya 2006; Doree 2007; Dunner 2007; Eisendrath 2016; El-Khalili 2010; Fang 2010/2011; Fava 1994a; Fava 2002; Fava 2012/Mischoulon 2012; Ferreri 2001; Fonagy 2015; GlaxoSmithKline 2009; Gulrez 2012; Joffe 1993; Joffe 2006; Kantor 1986; Katona 1995; Keitner 2009; Kennedy 2003; Kocsis 2009/Klein 2011; Lavretsky 2011; Lenox-Smith 2008; Licht 2002; Mahmoud 2007; Marcus 2008; McIntyre 2007; Mota-Pereira 2011; Nierenberg 2003a; Nierenberg 2006; Patkar 2006; Paykel 1999/Scott 2000; Peet 2002; Poirier 1999; Ravindran 2008a; Reeves 2008; Ruhe 2009; Rush 2006; Santos 2008; Schindler 2007; Schlogelhofer 2014; Schweizer 2001; Shelton 2005; Souery 2011a; Souza 2016; Stein 1993; Thase 2007; Town (unpublished); Trivedi 2006; Valenstein 2016; Watkins 2011a; Wiles 2013/2016; Zusky 1988). One hundred and forty-three studies were reviewed at full-text and excluded from this review. The most common reasons for exclusion were that there was non-randomised group assignment or not randomised at point of non-response, the intervention or comparison was not of interest (outside the protocol) or the sample size failed to meet our criteria of at least ten participants per arm (please note that an exception was made on the minimum sample size for lithium trials so as not to exclude a large proportion of the available evidence).

Studies not included in this review with reasons for their exclusions are provided in Appendix J5.

Meta-analyses were conducted according to further-line treatment strategy as follows:

- dose escalation strategies
- switching strategies (including switching to another antidepressant of the same class, switching to another antidepressant of a different class, and switching to a non-antidepressant treatment)
- augmentation strategies (including augmenting the antidepressant with another antidepressant, augmenting the antidepressant with a non-antidepressant agent and augmenting the antidepressant with a psychological intervention).

8.3.1 Dose escalation strategies

Evidence was found for three dose escalation treatment strategy comparisons as follows: increasing the dose of the antidepressant compared to continuing with the antidepressant at the same dose (see Table 93 for study characteristics); increasing the dose of the antidepressant compared to switching to another antidepressant (see Table 94 for study characteristics); increasing the dose of the antidepressant compared with augmenting with another antidepressant or non-antidepressant agent (see Table 95 for study characteristics).

Update 2017

Evidence for these comparisons are summarised in the clinical GRADE evidence profiles below (Table 96, Table 97 and Table 98). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J5.

Table 93: Study information table for trials included in the meta-analysis of increasing the dose of antidepressant versus continuing with the antidepressant at the same dose

	Increasing dose of SSRI versus continuing with SSRI at same dose
Total no. of studies (N randomised)	3 (430)
Study ID	Licht 2002[1] Ruhe 2009[2] Schweizer 2001[3]
Country	Denmark and Iceland1 Netherlands2 US3
Diagnostic status	DSM-IV MDD, without psychotic symptoms1 DSM-IV MDD, confirmed with SCID2 DSM-IV MDD[3]
Age range (mean)	Range NR (40.3)[1] Range NR (42.4)[2] Range NR (40.0)[3]
Sex (% female)	62[1] 67[2] 54[3]
Ethnicity (% BME)	NR[1,3] 40[2]
Mean age (SD) at first onset of depression	33 (12)[1] 37.6 (10.5)[2] NR[3]
Mean months (SD) since onset of current episode	Median: 4[1] NR[2] Mean NR (60% ≥12 months)[3]
No. (SD) of previous depressive episodes	Median: 2[1] 1.7 (1.4) [2] Mean NR (53% single episode)[3]
Details of inadequate response/treatment resistance	Inadequate response (<50% improvement on HAMD) to 6 weeks of open-label treatment with sertraline (50-100mg/day) [1] Inadequate response (<50% improvement on HAMD) to 6 weeks, open-label paroxetine treatment (20 mg/day) [2] Inadequate response (failure to achieve remission [HAMD-17>8]) to 3-week open-label prospective treatment phase with sertraline (50mg/day)[3]
Augmented/previous treatment	Previous treatment: Sertraline (100mg/day) [1] Previous treatment: Paroxetine (20mg/day) [2] Previous treatment: Sertraline (50mg/day)[3]
Baseline severity	NR[1] HAMD 20.6 (Less severe) [2]

Update 2017

	Increasing dose of SSRI versus continuing with SSRI at same dose
	HAMD 23.4 (Less severe) [3]
Intervention details (mean dose)	Sertraline (200mg/day; + placebo) [1] Paroxetine (30-50mg/day; mean dose 46.7mg/day) [2] Sertraline (150mg/day)[3]
Comparator details (mean dose, if applicable)	Sertraline (100mg/day; + placebo) [1] Paroxetine (20mg/day; + placebo) [2] Sertraline (50mg/day)[3]
Treatment length (weeks)	5[1,3] 6[2]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Licht 2002; [2]Ruhe 2009; [3]Schweizer 2001 Note that Licht[1] is a three-armed study and demographics reported here are for all three arms combined	

Table 94: Study information table for trials included in the meta-analysis of increasing the dose of antidepressant versus switching to another antidepressant

	Increasing dose of SSRI versus switch to SNRI
Total no. of studies (N randomised)	1 (484)
Study ID	Bose 2012
Country	US
Diagnostic status	DSM-IV MDD, confirmed with MINI
Age range (mean)	Range NR (42.3)
Sex (% female)	59
Ethnicity (% BME)	22
Mean age at first onset of depression	30.7 (SD NR)
Mean months since onset of current episode	11.1 (SD NR)
No. of previous depressive episodes	NR
Details of inadequate response/treatment resistance	Inadequate response (<50% improvement on MADRS) to 2 weeks of single-blind escitalopram (10mg/day)
Augmented/previous treatment	Previous treatment: Escitalopram (10mg/day)
Baseline severity	MADRS 34.8 (More severe)
Intervention details (mean dose)	Escitalopram (20mg/day)
Comparator details (mean dose, if applicable)	Duloxetine (60mg/day)
Treatment length (weeks)	8
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation	

Update 2017

Table 95: Study information table for trials included in the meta-analysis of increasing the dose of antidepressant versus augmenting with another antidepressant or non-antidepressant agent

	Increasing dose of SSRI versus TCA augmentation	Increasing dose of SSRI versus lithium augmentation	Increasing dose of SSRI versus TeCA augmentation
Total no. of studies (N randomised)	2 (142)	2 (142)	1 (295)
Study ID	Fava 1994a[1] Fava 2002[2]	Fava 1994a[1] Fava 2002[2]	Licht 2002
Country	US	US	Denmark and Iceland
Diagnostic status	DSM-III-R MDD	DSM-III-R MDD	DSM-IV MDD, without psychotic symptoms
Age range (mean)	18-65 (39.6)[1] Range NR (41.6)[2]	18-65 (39.6)[1] Range NR (41.6)[2]	Range NR (40.3)
Sex (% female)	61[1] 49[2]	61[1] 49[2]	62
Ethnicity (% BME)	NR	NR	NR
Mean age at first onset of depression	NR	NR	33 (12)
Mean months since onset of current episode	NR	NR	Median: 4
No. of previous depressive episodes	NR	NR	Median: 2
Details of inadequate response/treatment resistance	Inadequate response (defined as failure to achieve a 50% or greater reduction in HAMD score and a HAMD score of ≥10) to 8 weeks of open-label treatment with fluoxetine (20mg/day)	Inadequate response (defined as failure to achieve a 50% or greater reduction in HAMD score and a HAMD score of ≥10) to 8 weeks of open-label treatment with fluoxetine (20mg/day)	Inadequate response (<50% improvement on HAMD) to 6 weeks of open-label treatment with sertraline (50-100mg/day)
Augmented/previous treatment	Augmented/previous antidepressant: Fluoxetine (20mg/day)	Augmented/previous antidepressant: Fluoxetine (20mg/day)	Augmented/previous antidepressant: Sertraline (100mg/day)
Baseline severity	HAMD 14.5 (Less severe)[1] HAMD 16.6 (Less severe)[2]	HAMD 14.5 (Less severe)[1] HAMD 16.6 (Less severe)[2]	NR
Intervention details (mean dose)	Fluoxetine (40-60mg/day)	Fluoxetine (40-60mg/day)	Sertraline (200mg/day; + placebo)
Comparator details (mean dose, if applicable)	Desipramine (25-50mg/day, + fluoxetine 20mg/day)	Lithium (300-600mg/day, + fluoxetine 20mg/day)	Mianserin (10-30mg/day; + sertraline [100mg/day])
Treatment length (weeks)	4	4	5

Notes:

Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation

[1]Fava 1994a; [2]Fava 2002

Note that Fava 1994a[1], Fava 2002[2] and Licht 2002 are three-armed trials and demographics reported here are for all three arms combined

Update 2017

Table 96: Summary of findings table for increasing the dose of antidepressant versus continuing with the antidepressant at the same dose

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk: Continuing with the antidepressant at the same dose	Corresponding risk: Increasing the dose of antidepressant				
Remission Number of people scoring ≤7/8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 5-6 weeks	Study population **315 per 1000** **Moderate** **324 per 1000**	**337 per 1000** (198 to 576) **347 per 1000** (204 to 593)	**RR 1.07** (0.63 to 1.83)	327 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
Response Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 5-6 weeks	Study population **632 per 1000** **Moderate** **537 per 1000**	**506 per 1000** (411 to 626) **430 per 1000** (349 to 532)	**RR 0.8** (0.65 to 0.99)	252 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,4]	
Response Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: 5-6 weeks	**Study population** **778 per 1000** **Moderate** **712 per 1000**	**801 per 1000** (459 to 1000) **733 per 1000** (420 to 1000)	**RR 1.03** (0.59 to 1.8)	270 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3,5]	
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 6 weeks		The mean depression symptomatology in the intervention groups was **1.7 higher** (1.09 lower to 4.49 higher)		57 (1 study)	⊕⊖⊖⊖ **very low**3,6,7	
Discontinuation for any reason Number of participants	**Study population** **139 per 1000**	**97 per 1000** (29 to 330)	**RR 0.7** (0.21 to 2.38)	332 (3 studies)	⊕⊖⊖⊖ **very low**1,3,8,9	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant at the same dose	Increasing the dose of antidepressant				
discontinuing for any reason (including adverse events) Follow-up: 5-6 weeks	**Moderate** **135 per 1000**	**94 per 1000** (28 to 321)				
Discontinuation due to adverse events Number of participants discontinuing due to adverse events Follow-up: mean 6 weeks	**Study population** **133 per 1000** **Moderate** **133 per 1000**	**15 per 1000** (1 to 264) **15 per 1000** (1 to 263)	**RR 0.11** (0.01 to 1.98)	60 (1 study)	⊕⊖⊖⊖ **very low**3,9	

Notes:
[1] Unclear blinding of intervention administrator and outcome assessor
[2] 95% CI crosses line of no effect, and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] Data cannot be extracted/is not reported for all outcomes and study funded by pharmaceutical company
[4] Events<300
[5] I-squared>80%
[6] Unclear blinding of intervention administration and possible risk of attrition bias difference in drop-out between groups>20%) although ITT analysis used
[7] 95% CI crosses both line of no effect and threshold for clinically important harm (SMD 0.5)
[8] I-squared>50%
[9] 95% CI crosses line of no effect and both threshold for clinically important benefit (RR 0.75) and threshold for clinically important harm (RR 1.25)

Table 97: Summary of findings table for increasing the dose of antidepressant versus switching to another antidepressant

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Switching to another antidepressant	Increasing the dose of antidepressant				
Remission Number of people scoring ≤10 on Montgomery Asberg	**Study population** **420 per 1000**	**541 per 1000** (449 to 655)	**RR 1.29** (1.07 to 1.56)	472 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Switching to another antidepressant	**Increasing the dose of antidepressant**				
Depression Rating Scale (MADRS) Follow-up: mean 8 weeks	**Moderate** **420 per 1000**	**542 per 1000** (449 to 655)				
Response Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 8 weeks	**Study population** **700 per 1000** **Moderate** **700 per 1000**	**728 per 1000** (651 to 819) **728 per 1000** (651 to 819)	**RR 1.04** (0.93 to 1.17)	472 (1 study)	⊕⊕⊖⊖ **low**[1,3]	
Response Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 8 weeks	**Study population** **749 per 1000** **Moderate** **749 per 1000**	**771 per 1000** (697 to 854) **771 per 1000** (697 to 854)	**RR 1.03** (0.93 to 1.14)	472 (1 study)	⊕⊕⊖⊖ **low**[1,3]	
Depression symptomatology Quick Inventory of Depressive Symptomatology (QIDS; change score) Follow-up: mean 8 weeks		The mean depression symptomatology in the intervention groups was **0.9 lower** (1.88 lower to 0.08 higher)		472 (1 study)	⊕⊕⊖⊖ **low**[1,3]	
Discontinuation for any reason Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 8 weeks	**Study population** **215 per 1000** **Moderate** **215 per 1000**	**235 per 1000** (168 to 327) **234 per 1000** (168 to 327)	**RR 1.09** (0.78 to 1.52)	484 (1 study)	⊕⊕⊖⊖ **low**[4]	
Discontinuation due to adverse events Number of	**Study population** **53 per 1000**	**54 per 1000** (26 to 115)	**RR 1.03** (0.49 to 2.18)	484 (1 study)	⊕⊖⊖⊖ **very low**[3,5]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Switching to another antidepressant	Increasing the dose of antidepressant				
participants discontinuing due to adverse events Follow-up: mean 8 weeks	**Moderate** **53 per 1000**	**55 per 1000** (26 to 116)				

Notes:
[1] Unclear blinding of outcome assessment and risk of attrition bias (drop-out>20% [23%]) although difference between groups<20% and ITT analysis
[2] Events<300
[3] Data cannot be extracted/is not reported for all outcomes and study funded by pharmaceutical company
[4] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)
[5] 95% CI crosses line of no effect and both threshold for clinically important benefit (RR 0.75) and threshold for clinically important harm (RR 1.25)

Table 98: Summary of findings table for increasing the dose of antidepressant versus augmenting with another antidepressant or non-antidepressant agent

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmenting with another antidepressant/non-antidepressant agent	Increasing the dose of antidepressant				
Remission (increasing dose of SSRI versus TCA augmentation) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 4 weeks	**Study population** **283 per 1000** **Moderate** **272 per 1000**	**452 per 1000** (257 to 794) **435 per 1000** (248 to 764)	**RR 1.6** (0.91 to 2.81)	94 (2 studies)	⊕⊖⊖⊖ **very low**[1,2]	
Remission (increasing dose of SSRI versus lithium augmentation) Number of people	**Study population** **250 per 1000** **Moderate**	**458 per 1000** (257 to 812)	**RR 1.83** (1.03 to 3.25)	96 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,4]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmenting with another antidepressant/non-antidepressant agent	**Increasing the dose of antidepressant**				
scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 4 weeks	**261 per 1000**	**478 per 1000** (269 to 848)				
Remission (increasing dose of SSRI versus TeCA [mianserin] augmentation) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 4 weeks	**Study population**		**RR 0.66** (0.45 to 0.97)	195 (1 study)	⊕⊖⊖⊖ **very low**[2,4,5]	
	439 per 1000	**290 per 1000** (197 to 426)				
	Moderate					
	439 per 1000	**290 per 1000** (198 to 426)				
Response (increasing dose of SSRI versus TeCA [mianserin] augmentation) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 5 weeks	**Study population**		**RR 0.83** (0.66 to 1.03)	195 (1 study)	⊕⊖⊖⊖ **very low**[2,5,6]	
	673 per 1000	**559 per 1000** (444 to 694)				
	Moderate					
	674 per 1000	**559 per 1000** (445 to 694)				
Response (increasing dose of SSRI versus TeCA [mianserin] augmentation) Number of people rated as much or very much improved on Clinical Global	**Study population**		**RR 0.88** (0.74 to 1.04)	195 (1 study)	⊕⊖⊖⊖ **very low**[2,5,6]	
	776 per 1000	**682 per 1000** (574 to 807)				
	Moderate					
	776 per 1000	**683 per 1000** (574 to 807)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmenting with another antidepressant/non-antidepressant agent	**Increasing the dose of antidepressant**				
Impressions scale (CGI-I) Follow-up: mean 5 weeks						
Depression symptomatology (increasing dose of SSRI versus TCA augmentation) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 4 weeks		The mean depression symptomatology (increasing dose of SSRI versus TCA augmentation) in the intervention groups was **2.97 lower** (6.08 lower to 0.13 higher)		94 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,7,8]	
Depression symptomatology (increasing dose of SSRI versus lithium augmentation) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 4 weeks		The mean depression symptomatology (increasing dose of SSRI versus lithium augmentation) in the intervention groups was **2 lower** (4.32 lower to 0.33 higher)		96 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,8]	
Discontinuation for any reason (increasing dose of SSRI versus TCA augmentation) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 4 weeks	**Study population** **174 per 1000** **Moderate** **199 per 1000**	**101 per 1000** (37 to 285) **115 per 1000** (42 to 326)	**RR 0.58** (0.21 to 1.64)	94 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,9]	
	Study population					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmenting with another antidepressant/non-antidepressant agent	**Increasing the dose of antidepressant**				
Discontinuation for any reason (increasing dose of SSRI versus lithium augmentation) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 4 weeks	**146 per 1000**	**105 per 1000** (35 to 308)				
	Moderate					
	145 per 1000	**104 per 1000** (35 to 306)	**RR 0.72** (0.24 to 2.11)	96 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,9]	
Discontinuation for any reason (increasing dose of SSRI versus TeCA [mianserin] augmentation) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 5 weeks	**Study population**		**RR 0.88** (0.47 to 1.67)	196 (1 study)	⊕⊖⊖⊖ **very low**[2,9]	
	173 per 1000	**153 per 1000** (82 to 290)				
	Moderate					
	174 per 1000	**153 per 1000** (82 to 291)				
Discontinuation due to adverse events (increasing dose of SSRI versus TCA augmentation) Number of participants discontinuing due to adverse events Follow-up: mean 4 weeks	**Study population**		**RR 0.16** (0.01 to 3.09)	27 (1 study)	⊕⊖⊖⊖ **very low**[2,3,9]	
	167 per 1000	**27 per 1000** (2 to 515)				
	Moderate					
	167 per 1000	**27 per 1000** (2 to 516)				
Discontinuation due to adverse events (increasing dose	**Study population**		**RR 0.31** (0.01 to 7.09)	29 (1 study)	⊕⊖⊖⊖ **very low**[2,3,9]	
	71 per 1000	**22 per 1000** (1 to 506)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmenting with another antidepressant/non-antidepressant agent	Increasing the dose of antidepressant				
of SSRI versus lithium augmentation) Number of participants discontinuing due to adverse events Follow-up: mean 4 weeks	Moderate 71 per 1000	22 per 1000 (1 to 503)				

Notes:
[1] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[2] Data cannot be extracted/is not reported for all outcomes and/or funding from pharmaceutical company
[3] Unclear randomization method and allocation concealment and unclear blinding of intervention administration and outcome assessment
[4] Events<300
[5] Unclear blinding of intervention administration and outcome assessment
[6] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[7] I-squared>50%
[8] 95% CI crosses both line of effect and threshold for clinically important benefit (SMD -0.5)
[9] 95% CI crosses line of no effect and both threshold for clinically important benefit (RR 0.75) and threshold for clinically important harm (RR 1.25)

Update 2017

8.3.21 Augmentation strategies

Evidence was found for eight augmentation treatment strategy comparisons as follows: augmenting the antidepressant with another antidepressant or a non-antidepressant agent compared to augmentation with placebo (see Table 99, Table 100 and Table 101 for study characteristics); augmenting the antidepressant with another antidepressant or a non-antidepressant agent compared to continuing with the antidepressant-only (see Table 103, Table 104 and Table 105 for study characteristics); head-to-head comparisons of pharmacological augmentation agents (see Table 107, Table 109, Table 111, Table 113 and Table 115 for study characteristics); augmenting the antidepressant with a psychological intervention compared to augmentation with attention-placebo (see Table 117 for study characteristics); augmenting the antidepressant with a psychological intervention compared to continuing with the antidepressant-only (see Table 119, Table 120 and Table 121 for study characteristics); augmenting the antidepressant with a psychological intervention compared to augmenting the antidepressant with a non-antidepressant agent (see Table 123 for study characteristics); head-to-head comparisons of psychological augmentation interventions (see Table 125 for study characteristics); augmenting the antidepressant with a physical intervention compared to augmentation with attention-placebo (see Table 127 for study characteristics).

Evidence for these comparisons are summarised in the clinical GRADE evidence profiles below (see Table 102, Table 106, Table 108, Table 110, Table 112, Table 114, Table 116, Table 118, Table 122, Table 124, Table 126 and Table 128). See also the full GRADE

evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J5.

Table 99: Study information table for trials included in the meta-analysis of augmenting the antidepressant with another antidepressant or a non-antidepressant agent versus placebo (part 1)

	Atypical antidepressant	Antipsychotic	Lithium
Total no. of studies (N randomised)	2 (86)	10 (2709)	8 (260)
Study ID	Carpenter 2002[1] Gulrez 2012[2]	Bauer 2009[3] Berman 2007[4] Berman 2009[5] El-Khalili 2010[6] Fava 2012/Mischoulon 2012[7] Keitner 2009[8] Mahmoud 2007[9] Marcus 2008[10] McIntyre 2007[11] Reeves 2008[12]	Browne 1990[13] Joffe 1993[14] Joffe 2006[15] Kantor 1986[16] Katona 1995[17] Nierenberg 2003a[18] Stein 1993[19] Zusky1988[20]
Country	US[1] India[2]	Australia, Canada, Europe and South Africa[3] US[4,5,7,8,9,10,12] US and Sweden[6] Canada[11]	Canada[13,14,15,16] UK[17,19] US[18,20]
Diagnostic status	DSM-IV 88.5% unipolar MDD (recurrent) and 11.5% bipolar II disorder (current episode depressed) [1] DSM–IV-TR MDD[2]	DSM-IV-TR MDD[3,4,5,10] DSM-IV MDD[6,8,9,11] SCID for DSM Disorders major depressive episode (MDE) diagnosis deemed 'valid' using the SAFER criteria interview[7] DSM-IV MDD, currently experiencing a depressive episode with suicidal ideation[12]	DSM-III 82% unipolar MDD and 18% bipolar[13] RDC criteria for unipolar, nonpsychotic MDD[14] DSM-IV criteria for nonpsychotic, unipolar MDD[15] Unipolar MDD[16] DSM-III MDD or bipolar disorder[17] DSM-III-R MDD[18] RDC MDD[19] DSM-III MDD, without psychosis[20]
Age range (mean)	Range NR (46.3) [1] 18-75 (41.2) [2]	18-65 (45.4)[3,4,5] 18-65 (45.5)[6] 18-65 (45)[7] 20-63 (45.2)[8] 20-65 (46.1)[9] 18-65 (44.5)[10] Range NR (44.5)[11] 19-60 (44.0)[12]	26-66 (42.7)[13] Range NR (37.4) [14] 23-52 (39.2) [15] NR[16] Range NR (40.0) [17] Range NR (38.4) [18] Range NR (47.2) [19] 18-80 (45.8) [20]
Sex (% female)	62[1] 52[2]	68[3,7] 63[4] 73[5]	59[13] 61[14] 83[15]

Update 2017

	Atypical antidepressant	Antipsychotic	Lithium
		72[6] 59[8] 74[9] 67[10] 62[11] 70[12]	NR[16] 56[17] 46[18] 79[19] 81[20]
Ethnicity (% BME)	NR	2[3] 10[4,6,8] 13[5] 19[7] 24[9] 11[10] NR[11,12]	NR
Mean age (SD) at first onset of depression	NR	NR[3,4,5,6,8,9,10,11,12] 16.8 (13.6)[7]	NR[13,14,15,16,17,19,20] 19.9 (11.5)[18]
Mean months (SD) since onset of current episode	6.4 (5.3) [1] NR[2]	NR[3,6,7,9,11,12] 41.1 (56.5)[4] 18 (SD NR)[5] 44.4 (70.2)[8] 46.1 (79)[10]	48.5 (SD NR)[13] NR[14,15,16,17,19.20] 91.1 (102.6)[18]
No. (SD) of previous depressive episodes	2.4 (1.7) [1] NR[2]	NR[3,4,6,7,9,11,12] 5.8 (9.1)[5] 3.8 (1.5)[8] 6.8 (13.6)[10]	NR[13,14,15,16,17,19,20] 0.6 (1.0)[18]
Details of inadequate response/treatment resistance	Inadequate response (HAMD total score>12) after at least 4 weeks of standard antidepressant monotherapy at maximum recommended or tolerated doses[1] Inadequate response (HAMD score ≥16) after 4 weeks of SSRI treatment[2]	Inadequate response to at least 1 previous course of antidepressants at adequate dose for at least 3[12]/5[8]/6[3,6,11]/8[7] weeks Inadequate response to a prospective 4-week treatment phase[9] TRD: Inadequate response to at least 1 previous course of antidepressants at an adequate dose for at least 6 weeks (for the current episode) and failure to respond to a prospective 8-week treatment phase[4,5,10]	Inadequate response to at least 1 previous course of antidepressants at adequate dose for at least 3[16,19]/4[13,20]/5[14,15]/6[17] weeks TRD: Inadequate response to at least 1 previous course of antidepressants (for the current episode) and failure to respond to a prospective 6-week treatment phase[18]
Augmented/previous treatment	Augmented AD: 85% SSRIs (31% sertraline [100-200 mg/day]; 19% citalopram [30-60 mg/day]; 19% fluoxetine [40-50 mg/day]; 12% paroxetine [30-40 mg/day]; 4%	Augmented antidepressant: Predominantly SSRIs or venlafaxine Averaged across the 9 studies reporting proportions: Escitalopram 22%; Sertraline 16%;	Augmented antidepressant: 29% imipramine; 24% doxepin; 12% maprotiline; 12% trimipramine; 12% clomipramine; 6% amitriptyline; 6% desipramine[13]

Update 2017

	Atypical antidepressant	Antipsychotic	Lithium
	fluvoxamine [300 mg/day]); 12% venlafaxine (200-300 mg/day); 4% bupropion (450 mg/day)[1] Augmented AD: SSRI: 40% sertraline (mean dose 106mg); 37% escitalopram (mean dose 21mg); 13% citalopram (mean dose 28mg); 10% paroxetine (mean dose 33mg)[2]	Fluoxetine 13%; Paroxetine 8%; Citalopram 7% Venlafaxine 21% Bupropion 5% Duloxetine 4% Amitriptyline 1% Desvenlafaxine 1% Other 1%	90% desipramine; 10% imipramine (mean dose of desipramine or imipramine 201mg/day [range 150-300mg/day])[14] 78% SSRI[15] 29% amtriptyline (200mg/day); 29% imipramine (150 or 250mg/day); 29% doxepin (100 or 150mg/day); 14% amoxapine (250mg/day)[16] Fluoxetine (20mg/day) or lofepramine (140-210mg/day)[17] Nortriptyline (mean dose 116.7mg [SD=31.6])[18] 50% amitriptyline; 18% dothiepin; 12% trimipramine; 6% imipramine; 6% doxepin; 3% clomipramine; 3% lofepramine; 3% protriptyline. Mean TCA dose at baseline 161.7mg/day (SD=62.5)[19] 31% desipramine; 13% amitriptyline; 13% trazodone; 13% imipramine; 13% nortriptyline; 6% maprotiline; 6% doxepin; 6% phenelzine[20]
Baseline severity	HAMD 22.3 (Less severe)[1] MADRS 20.5 (Less severe)[2]	HAMD 24.6 (More severe)[3,9] MADRS 26 (Less severe)[4] MADRS 26.9 (Less severe)[5] HAMD 24.1 (More severe)[6] MADRS 31.1 (More severe)[7] MADRS 25.7 (Less severe)[8] MADRS 26.1 (Less severe)[10]	HAMD 23.4 (Less severe)[13] HAMD 19.5 (Less severe)[14,15] HAMD 23.3 (Less severe)[16] HAMD 18.6 (Less severe)[17] NR[18] MADRS 29.9 (More severe)[19] HAMD 22.6 (Less severe)[20]

Update 2017

	Atypical antidepressant	Antipsychotic	Lithium
		HAMD 23.3 (Less severe)[11] MADRS 35.5 (More severe)[12]	
Intervention details (mean dose)	Mirtazapine (final dose: 31% 15mg/69% 30mg) [1] Bupropion Sustained Release (150-300mg/day) [2]	Quetiapine extended-release (two dose arms combined: 150mg/day and 300mg/day)[3,6] Aripiprazole (2-20mg/day); mean final dose 11.8mg/day[4]; mean final dose 10.7mg/day[5] Aripiprazole low dose (2mg/day)[7] Risperidone (0.5-3mg/day; mean final dose 1.6 mg/day)[8] Risperidone (0.25-2mg/day)[9;]; mean final dose 1.2mg/day[12] Aripiprazole (5-20mg/day; mean final dose 11mg/day)[10] Quetiapine (50-600mg/day; mean dose 182mg/day)[11]	Lithium 900mg/day[13] Lithium 900-1200mg/day (target plasma level 0.55 nmol/L; mean dose 935.3mg/day)[14] Lithium 600-900mg/day[15] Lithium 900mg/day[16] Lithium 400-800mg/day (target plasma level 0.6-1.0 mmol/l)[17] Lithium (no further detail reported)[18] Lithium 250mg/day (+2 placebo tablets)[19] Lithium 300-900mg/day[20]
Comparator details (mean dose, if applicable)	Placebo	Placebo[3,6,7,8,9,11] Placebo (2-20mg/day); mean final dose 15.7mg/day[4]; mean final dose 13.9mg/day[5] Placebo (5-20mg/day; mean final dose 15.3mg/day)[10] Placebo (0.25-2mg/day; mean final dose 1.5mg/day)[12]	Placebo[13,15,17,18] Placebo 900-1200mg/day[14] Placebo 3 capsules/day[16,19] Placebo 1-3 capsules/day[20]
Treatment length (weeks)	4	6[3,4,5,6,9,10] 4[7,8] 8[11,12]	0.3[13,16] 2[14,15] 6[17,18] 3[19,20]

Notes:

Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression

[1]Carpenter 2002; [2]Gulrez 2012; [3]Bauer 2009; [4]Berman 2007; [5]Berman 2009; [6]El-Khalili 2010; [7]Fava 2012/Mischoulon 2012; [8]Keitner 2009; [9]Mahmoud 2007; [10]Marcus 2008; [11]McIntyre 2007; [12]Reeves 2008; [13]Browne 1990; [14]Joffe 1993; [15]Joffe 2006; [16]Kantor 1986; [17]Katona 1995; [18]Nierenberg 2003a; [19]Stein 1993; [20]Zusky1988

Note that Bauer 2009[3], El-Khalili 2010[6], Fava 2012/Mischoulon 2012[7] and Joffe 1993[14] are three-armed trials and demographics reported here are for all three arms combined, and Joffe 2006[15] is a four-armed trial and demographics reported here are for all four arms combined

Update 2017

Table 100: Study information table for trials included in the meta-analysis of augmenting the antidepressant with another antidepressant or a non-antidepressant agent versus placebo (part 2)

	Thyroid hormone	Anticonvulsant	Stimulant
Total no. of studies (N randomised)	2 (69)	2 (130)	2 (205)
Study ID	Joffe 1993[1] Joffe 2006[2]	Barbee 2011[3] Santos 2008[4]	Patkar 2006[5] Ravindran 2008a[6]
Country	Canada	US[3] Brazil[4]	US[5] Canada[6]
Diagnostic status	RDC criteria for unipolar, nonpsychotic MDD[1] DSM-IV criteria for nonpsychotic, unipolar MDD[2]	DSM-IV/ICD-10 unipolar MDD, confirmed by the MINI[3] DSM-IV MDD (single or recurrent)[4]	DSM-IV MDD, without psychotic features, confirmed with MINI[5] DSM-IV-TR MDD, without psychotic features, confirmed by MINI[6]
Age range (mean)	Range NR (37.4)[1] 23-52 (39.2)[2]	18-65 (45.2)[3] Range NR (27.5)[4]	Range NR (48.5)[5] Range NR (43.8)[6]
Sex (% female)	61[1] 83[2]	69[3] 74[4]	63[5] 65[6]
Ethnicity (% BME)	NR	NR	40[5] 2[6]
Mean age (SD) at first onset of depression	NR	26.2 (13.4)[3] 28.5 (12.7)[4]	27.8 (14.5)[5] NR[6]
Mean months (SD) since onset of current episode	NR	26.9 (36.9)[3] 32.3 (49.9)[4]	19.4 (23.4)[1] 21.8 (47.5)[2]
No. (SD) of previous depressive episodes	NR	9.2 (20.4)[3] 6.5 (6.8)[4]	NR
Details of inadequate response/treatment resistance	Inadequate response (had a score≥16 on the 17-item HAMD) to a previous adequate trial of desipramine hydrochloride (90%) or imipramine hydrochloride (10%) ≥5 weeks (at a minimum dose of 2.5mg/kg of body weight per day)[1] Inadequate response to a trial of antidepressants at usual dosages (moclobemide 600 to 750 mgdaily, nefazodone 150 to 300 mg daily, paroxetine 20 to 60 mg daily, sertraline 100 to 200 mg daily, fluoxetine 30 to 40 mg daily,	TRD: Inadequate response to ≥1 previous 6-week antidepressant treatment for current episode, and failure to respond to open-label prospective 8-week treatment with paroxetine[3] TRD: Inadequate response to treatment with ≥ 2 antidepressants of different classes at the maximum-tolerated dose for ≥6 weeks[4]	Inadequate response to ≥1 antidepressant at study entry, at an acceptable therapeutic dose for ≥6 weeks. 70% had failed multiple antidepressant trials for the current MDD episode[5] Inadequate response to 1-3 previous antidepressant monotherapies (including current antidepressant) of adequate dose and duration and at entry were taking an adequate dose of an antidepressant during the current depressive episode ≥4 weeks[6]

Update 2017

	Thyroid hormone	Anticonvulsant	Stimulant
	fluvoxamine 150 to 300 mg daily, and venlafaxine 187.5 to 375 mg daily) for at least 5 weeks[2]		
Augmented/previous treatment	Augmented antidepressant: 90% desipramine; 10% imipramine (mean dose of desipramine or imipramine 201mg/day [range 150-300mg/day]) [1] Augmented antidepressant: 78% SSRI[2]	Augmented antidepressant: Paroxetine (mean 44.84mg/day) or paroxetine CR (mean 49.53mg/day) [3] Augmented antidepressant: 29% SSRI; 21% TCA; 21% venlafaxine; 9% bupropion; 9% milnacipran; 12% other[4]	Augmented antidepressant: NR (pre-existing antidepressant dose was unchanged)
Baseline severity	HAMD 19.5 (Less severe)	MADRS 27 (More severe)[3] MADRS 30.4 (More severe)[4]	HAMD 19.4 (Less severe)[5] MADRS 26.7 (Less severe)[6]
Intervention details (mean dose)	Liothyronine sodium (triiodothyronine, T3) 37.5µg[1] Triiodothyronine (T3) 37.5 µg[2]	Lamotrigine (25-400mg/day; mean final dose 271.88 mg/day) [3] Lamotrigine (50-200mg/day) [4]	Methylphenidate extended release formulation (18-54mg/day); mean dose 34.2mg/day[5]; mean final dose 36.4mg/day[6]
Comparator details (mean dose, if applicable)	Placebo	Placebo	Placebo
Treatment length (weeks)	2	10[3] 8[4]	4[5] 5[6]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression [1]Joffe 1993; [2]Joffe 2006; [3]Barbee 2011; [4]Santos 2008; [5]Patkar 2006; [6]Ravindran 2008a Note that Joffe 1993[14] and Joffe 2006[15] are three-armed or four-armed trials respectively and demographics reported here are for all three/four arms combined			

Table 101: Study information table for trials included in the meta-analysis of augmenting the antidepressant with another antidepressant or a non-antidepressant agent versus placebo (part 3)

	Anxiolytic	Omega-3 fatty acid
Total no. of studies (N randomised)	1 (113)	1 (70)
Study ID	Appelberg 2001	Peet 2002
Country	Finland	UK
Diagnostic status	DSM-IV major depressive episode	Depression symptoms (HAMD score ≥15)
Age range (mean)	18-74 (44)	18-70 (44.7)
Sex (% female)	63	84
Ethnicity (% BME)	NR	NR

Update 2017

	Anxiolytic	Omega-3 fatty acid
Mean age (SD) at first onset of depression	NR	NR
Mean months (SD) since onset of current episode	30 (SD NR)	NR
No. (SD) of previous depressive episodes	NR	NR
Details of inadequate response/treatment resistance	Inadequate response (as judged by the psychiatrist in charge of treatment) to ≥ 6 weeks of treatment with fluoxetine (at a dose of ≥30mg/day for ≥4 weeks prior to inclusion) or citalopram (at a dose of ≥40mg/day for ≥4 weeks prior to inclusion)	Inadequate response (HAMD≥15) to ongoing treatment with antidepressant at an adequate dose
Augmented/previous treatment	Augmented antidepressants: 54% citalopram (40.3mg/day); 46% fluoxetine (34.7mg/day). Mean treatment time with an SSRI = 1.2 years	Augmented antidepressant: 71% SSRIs; 20% TCAs; 9% other
Baseline severity	NR	MADRS 22.7 (Less severe)
Intervention details (mean dose)	Buspirone (10-60mg/day; mean final dose 47mg/day)	Ethyl-eicosapemtaenoate (combined 3 dose groups: 1g/day, 2g/day and 4g/day)
Comparator details (mean dose, if applicable)	Placebo	Placebo
Treatment length (weeks)	6	12
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation		

Update 2017

Table 102: Summary of findings table for augmenting the antidepressant with another antidepressant or a non-antidepressant agent versus 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 Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
Remission (atypical antidepressant) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D)	**Study population**		**RR 2.72** (1.44 to 5.14)	86 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	200 per 1000	**544 per 1000** (288 to 1000)				
	Moderate					

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
Follow-up: mean 4 weeks	**183 per 1000**	**498 per 1000** (264 to 941)				
Remission (antipsychotic) Number of people scoring <10/11 on Montgomery Asberg Depression Rating Scale (MADRS)/≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 4-8 weeks	**Study population**		**RR 1.56** (1.36 to 1.78)	2581 (9 studies)	⊕⊕⊖⊖ **low**[3,4]	
	193 per 1000	**301 per 1000** (262 to 343)				
	Moderate					
	172 per 1000	**268 per 1000** (234 to 306)				
Remission (lithium) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 2-6 weeks	**Study population**		**RR 2.07** (1.16 to 3.69)	110 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	214 per 1000	**444 per 1000** (249 to 791)				
	Moderate					
	250 per 1000	**518 per 1000** (290 to 923)				
Remission (thyroid hormone [T3]) Number of people scoring <7 on Hamilton Rating Scale for Depression (HAM-D) AND responding (≥50% improvement on HAM-D) Follow-up: mean 2 weeks	**Study population**		**RR 3.29** (0.8 to 13.57)	33 (1 study)	⊕⊕⊖⊖ **low**[1,5]	
	125 per 1000	**411 per 1000** (100 to 1000)				
	Moderate					
	125 per 1000	**411 per 1000** (100 to 1000)				
Remission (stimulant [methylphenidate]) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D)	**Study population**		**RR 4** (0.47 to 33.73)	60 (1 study)	⊕⊖⊖⊖ **very low**[1,3,6]	
	33 per 1000	**133 per 1000** (16 to 1000)				
	Moderate					

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
Follow-up: mean 4 weeks	**33 per 1000**	**132 per 1000** (16 to 1000)				
Response (any augmentation agent) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 0.3-12 weeks	**Study population**		**RR 1.35** (1.23 to 1.49)	3110 (20 studies)	⊕⊕⊖⊖ **low**[3,4]	
	293 per 1000	**395 per 1000** (360 to 436)				
	Moderate					
	253 per 1000	**342 per 1000** (311 to 377)				
Response (atypical antidepressant) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 4 weeks	**Study population**		**RR 3.18** (1.05 to 9.62)	26 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	200 per 1000	**636 per 1000** (210 to 1000)				
	Moderate					
	200 per 1000	**636 per 1000** (210 to 1000)				
Response (antipsychotic) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 4-8 weeks	**Study population**		**RR 1.4** (1.25 to 1.57)	2604 (10 studies)	⊕⊕⊖⊖ **low**[3,4]	
	291 per 1000	**407 per 1000** (363 to 456)				
	Moderate					
	285 per 1000	**399 per 1000** (356 to 447)				
Response (lithium) Number of people showing ≥50% improvement on	**Study population**		**RR 1.55** (0.61 to 3.91)	76 (4 studies)	⊕⊖⊖⊖ **very low**[3,4,6]	
	158 per 1000	**245 per 1000** (96 to 617)				

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
Hamilton Rating Scale for Depression (HAM-D) Follow-up: 0.3-6 weeks	**Moderate**					
	151 per 1000	**234 per 1000** (92 to 590)				
Response (anticonvulsant [lamotrigine]) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-10 weeks	**Study population**		**RR 0.96** (0.59 to 1.56)	130 (2 studies)	⊕⊖⊖⊖ **very low**[3,6,7]	
	338 per 1000	**325 per 1000** (200 to 528)				
	Moderate					
	343 per 1000	**329 per 1000** (202 to 535)				
Response (omega-3 fatty acid) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 12 weeks	**Study population**		**RR 1.31** (0.51 to 3.38)	69 (1 study)	⊕⊖⊖⊖ **very low**[3,6,8]	
	235 per 1000	**308 per 1000** (120 to 795)				
	Moderate					
	235 per 1000	**308 per 1000** (120 to 794)				
Response (stimulant [methylphenidate]) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 4-5 weeks	**Study population**		**RR 1.21** (0.87 to 1.68)	205 (2 studies)	⊕⊖⊖⊖ **very low**[3,5,7]	
	363 per 1000	**439 per 1000** (316 to 609)				
	Moderate					
	325 per 1000	**393 per 1000** (283 to 546)				
Response (any augmentation agent) Number of people rated as much or very	**Study population**		**RR 1.29** (0.85 to 1.97)	257 (5 studies)	⊕⊖⊖⊖ **very low**[3,4,5]	
	285 per 1000	**367 per 1000** (242 to 561)				

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
much improved on Clinical Global Impressions scale (CGI-I) Follow-up: 4-8 weeks	**Moderate**					
	267 per 1000	**344 per 1000** (227 to 526)				
Response (atypical antidepressant) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 4 weeks	**Study population**		**RR 3.18** (1.05 to 9.62)	26 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	200 per 1000	**636 per 1000** (210 to 1000)				
	Moderate					
	200 per 1000	**636 per 1000** (210 to 1000)				
Response (lithium) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 6 weeks	**Study population**		**RR 1.18** (0.38 to 3.67)	35 (1 study)	⊕⊖⊖⊖ **very low**[1,3,6]	
	235 per 1000	**278 per 1000** (89 to 864)				
	Moderate					
	235 per 1000	**277 per 1000** (89 to 862)				
Response (anticonvulsant [lamotrigine]) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 8 weeks	**Study population**		**RR 0.67** (0.23 to 1.95)	34 (1 study)	⊕⊖⊖⊖ **very low**[6,9]	
	353 per 1000	**236 per 1000** (81 to 688)				
	Moderate					
	353 per 1000	**237 per 1000** (81 to 688)				
Response (anxiolytic) Number of people rated as much or very much improved on Clinical Global Impressions scale	**Study population**		**RR 1.06** (0.61 to 1.86)	102 (1 study)	⊕⊖⊖⊖ **very low**[3,6,10]	
	314 per 1000	**333 per 1000** (191 to 584)				
	Moderate					

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
(CGI-I) Follow-up: mean 6 weeks	**314 per 1000**	**333 per 1000** (192 to 584)				
Response (stimulant [methylphenidate]) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 4 weeks	**Study population**		**RR 1.62** (0.79 to 3.34)	60 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	
	267 per 1000	**432 per 1000** (211 to 891)				
	Moderate					
	267 per 1000	**433 per 1000** (211 to 892)				
Depression symptomatology (atypical antidepressant) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 4 weeks		The mean depression symptomatology (atypical antidepressant) in the intervention groups was **1.12 standard deviations lower** (1.96 to 0.27 lower)		26 (1 study)	⊕⊖⊖⊖ **very low**[1,3,11]	SMD -1.12 (-1.96 to -0.27)
Depression symptomatology (antipsychotic) Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 4-8 weeks		The mean depression symptomatology (antipsychotic) in the intervention groups was **0.4 standard deviations lower** (0.86 lower to 0.06 higher)		462 (3 studies)	⊕⊖⊖⊖ **very low**[3,4,12,13]	SMD -0.4 (-0.86 to 0.06)
Depression symptomatology (lithium) Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery Asberg Depression		The mean depression symptomatology (lithium) in the intervention groups was **0.23 standard deviations lower**		83 (3 studies)	⊕⊖⊖⊖ **very low**[3,4,13]	SMD -0.23 (-0.86 to 0.39)

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
Rating Scale (MADRS; change score) Follow-up: 2-3 weeks		(0.86 lower to 0.39 higher)				
Depression symptomatology (thyroid hormone [T3]) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 2 weeks		The mean depression symptomatology (thyroid hormone [t3]) in the intervention groups was **0.78 standard deviations lower** (1.5 to 0.07 lower)		33 (1 study)	⊕⊕⊖⊖ **low**[11,14]	SMD -0.78 (-1.5 to -0.07)
Depression symptomatology (anticonvulsant [lamotrigine]) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 8-10 weeks		The mean depression symptomatology (anticonvulsant [lamotrigine]) in the intervention groups was **0.13 standard deviations lower** (0.54 lower to 0.27 higher)		130 (2 studies)	⊕⊖⊖⊖ **very low**[3,13,15]	SMD -0.13 (-0.54 to 0.27)
Depression symptomatology (stimulant [methylphenidate]) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: mean 5 weeks		The mean depression symptomatology (stimulant [methylphenidate]) in the intervention groups was **0.06 standard deviations higher** (0.27 lower to 0.38 higher)		144 (1 study)	⊕⊖⊖⊖ **very low**[3,11,16]	SMD 0.06 (-0.27 to 0.38)
Discontinuation for any reason (atypical antidepressant) Number of participants discontinuing for any reason (including	**Study population**		**RR 0.68** (0.07 to 6.61)	86 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,17]	
	44 per 1000	**30 per 1000** (3 to 294)				
	Moderate					

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
adverse events) Follow-up: mean 4 weeks	**67 per 1000**	**46 per 1000** (5 to 443)				
Discontinuation for any reason (antipsychotic) Number of participants discontinuing for any reason (including adverse events) Follow-up: 4-8 weeks	**Study population**		**RR 1.24** (1.02 to 1.52)	2706 (10 studies)	⊕⊕⊖⊖ **low**[3,18]	
	135 per 1000	**167 per 1000** (137 to 205)				
	Moderate					
	141 per 1000	**175 per 1000** (144 to 214)				
Discontinuation for any reason (lithium) Number of participants discontinuing for any reason (including adverse events) Follow-up: 2-6 weeks	**Study population**		**RR 0.87** (0.41 to 1.84)	200 (6 studies)	⊕⊖⊖⊖ **very low**[1,3,17]	
	119 per 1000	**103 per 1000** (49 to 219)				
	Moderate					
	56 per 1000	**49 per 1000** (23 to 103)				
Discontinuation for any reason (thyroid hormone [T3]) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 2 weeks	See comment	See comment	Not estimable	51 (2 studies)	⊕⊕⊖⊖ **low**[2,14]	
Discontinuation for any reason (anticonvulsant [lamotrigine]) Number of participants discontinuing for any reason (including adverse events) Follow-up: 8-10 weeks	**Study population**		**RR 0.81** (0.48 to 1.38)	130 (2 studies)	⊕⊖⊖⊖ **very low**[3,17,19]	
	323 per 1000	**262 per 1000** (155 to 446)				
	Moderate					
	295 per 1000	**239 per 1000** (142 to 407)				

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
Discontinuation for any reason (anxiolytic) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 6 weeks	**Study population**		**RR 0.6** (0.24 to 1.53)	102 (1 study)	⊕⊖⊖⊖ **very low**[3,17,20]	
	196 per 1000	**118 per 1000** (47 to 300)				
	Moderate					
	196 per 1000	**118 per 1000** (47 to 300)				
Discontinuation for any reason (omega-3 fatty acid) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 12 weeks	**Study population**		**RR 0.52** (0.17 to 1.63)	70 (1 study)	⊕⊖⊖⊖ **very low**[3,17]	
	222 per 1000	**116 per 1000** (38 to 362)				
	Moderate					
	222 per 1000	**115 per 1000** (38 to 362)				
Discontinuation for any reason (stimulant [methylphenidate]) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 5 weeks	**Study population**		**RR 2.71** (0.91 to 8.12)	145 (1 study)	⊕⊖⊖⊖ **very low**[3,16,21]	
	56 per 1000	**151 per 1000** (51 to 451)				
	Moderate					
	56 per 1000	**152 per 1000** (51 to 455)				
Discontinuation due to adverse events (atypical antidepressant) Number of participants discontinuing due to adverse events Follow-up: mean 4 weeks	See comment	See comment	Not estimable	60 (1 study)	⊕⊖⊖⊖ **very low**[2,3,22]	
Discontinuation due to adverse events (antipsychotic) Number of participants	**Study population**		**RR 3.16** (1.97 to 5.06)	2706 (10 studies)	⊕⊖⊖⊖ **very low**[2,3,18]	
	17 per 1000	**54 per 1000** (34 to 87)				

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
discontinuing due to adverse events Follow-up: 4-8 weeks	**Moderate**					
	20 per 1000	**63 per 1000** (39 to 101)				
Discontinuation due to adverse events (lithium) Number of participants discontinuing due to adverse events Follow-up: 2-6 weeks	**Study population**		**RR 1.3** (0.33 to 5.14)	165 (5 studies)	⊕⊖⊖⊖ **very low**[3,14,17]	
	36 per 1000	**46 per 1000** (12 to 184)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				
Discontinuation due to adverse events (thyroid hormone [T3]) Number of participants discontinuing due to adverse events Follow-up: mean 2 weeks	See comment	See comment	Not estimable	51 (2 studies)	⊕⊕⊖⊖ **low**[2,14]	
Discontinuation due to adverse events (anticonvulsant [lamotrigine]) Number of participants discontinuing due to adverse events Follow-up: 8-10 weeks	**Study population**		**RR 1.12** (0.21 to 5.94)	130 (2 studies)	⊕⊖⊖⊖ **very low**[3,17,19]	
	154 per 1000	**172 per 1000** (32 to 914)				
	Moderate					
	104 per 1000	**116 per 1000** (22 to 618)				
Discontinuation due to adverse events (anxiolytic) Number of participants discontinuing due to adverse events Follow-up: mean 6 weeks	See comment	See comment	Not estimable	102 (1 study)	⊕⊖⊖⊖ **very low**[2,3,20]	

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				
Discontinuation due to adverse events (omega-3 fatty acid) Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Study population**		**RR 0.35** (0.02 to 5.25)	70 (1 study)	⊕⊖⊖⊖ **very low**[3,17]	
	56 per 1000	**19 per 1000** (1 to 292)				
	Moderate					
	56 per 1000	**20 per 1000** (1 to 294)				
Discontinuation due to adverse events (stimulant [methylphenidate]) Number of participants discontinuing due to adverse events Follow-up: 4-5 weeks	**Study population**		**RR 2.92** (0.21 to 40.65)	205 (2 studies)	⊕⊖⊖⊖ **very low**[3,12,17,18]	
	20 per 1000	**57 per 1000** (4 to 797)				
	Moderate					
	33 per 1000	**96 per 1000** (7 to 1000)				

Notes:
[1] Unclear randomisation method and method for allocation concealment. Blinding of intervention administration and outcome assessment is also unclear
[2] Events<300
[3] Data cannot be extracted/is not reported for all outcomes and/or funding from pharmaceutical company
[4] Unclear randomisation method and method for allocation concealment. Blinding of intervention administration and outcome assessment is also unclear for studies that make up >50% weighting in the analysis
[5] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[6] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[7] Significant group differences in baseline demographics at baseline in studies contributing to>50% weighting in analysis and unclear blinding of intervention administration and outcome assessment
[8] Unclear blinding of outcome assessment
[9] Unclear blinding of outcome assessment and unclear risk of attrition bias (drop-out>20% [21%] but difference between groups<20% and ITT analysis)
[10] Unclear randomisation method and method of allocation concealment. Blinding of outcome assessment is also unclear
[11] N<400
[12] I-squared>50%
[13] 95% CI crosses both line of no effect and threshold for clinically important benefit (SMD -0.5)
[14] Unclear randomisation method and method of allocation concealment, and blinding of intervention administration unclear

Update 2017

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	Augmenting the antidepressant with another antidepressant or a non-antidepressant agent				

[15] High risk of bias associated with randomisation method as significant differences between groups at baseline in studies contributing >50% to weighting in analysis. Unclear blinding of outcome assessment and unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[16] High risk of bias associated with randomisation method as significant difference between groups at baseline and method of allocation concealment is unclear. Blinding of intervention administration is also unclear
[17] 95% CI crosses line of no effect and both threshold for clinically important benefit (RR 0.75) and threshold for clinically important harm (RR 1.25)
[18] Unclear or high risk of bias associated with randomisation method, unclear method of allocation concealment, and unclear blinding of intervention administration in studies contributing to >50% of weighting in analysis
[19] High risk of bias associated with randomisation method as significant difference between groups at baseline
[20] Unclear randomisation method and method of allocation concealment
[21] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)
[22] Unclear randomisation method and method of allocation concealment, and blinding of intervention administration is unclear

Update 2017

Sub-analyses of the antipsychotic augmentation versus placebo comparison were performed, see forest plots in Appendix M, comparing non-sedating (aripiprazole) and sedating (quetiapine or risperidone) antipsychotics, in order to explore whether sedative effects might account for some of the apparent therapeutic benefits. However, the results of this analysis are inconclusive. The test for subgroup differences is non-significant for the rate of remission ($Chi^2 = 0.80$, df = 1, p = 0.37), discontinuation for any reason ($Chi^2 = 0.01$, df = 1 (P = 0.92) and discontinuation due to adverse events ($Chi^2 = 0.54$, df = 1, p = 0.46). For depression symptomatology, the test for subgroup differences is statistically significant ($Chi^2 = 8.15$, df = 1, p = 0.004) and suggests clinically important and statistically significant benefits of antipsychotic augmentation for sedating antipsychotics (K=2; N=241; SMD -0.64 [-0.90, -0.38]) but not for non-sedating antipsychotics (K=1; N=221; SMD -0.06 [-0.36, 0.25]). However, conversely, there is a trend for a statistically significant subgroup difference ($Chi^2 = 3.53$, df = 1, p = 0.06), for the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or MADRS), but here the benefits for both sedating (K=6; N= 1313; RR 1.29 [1.14, 1.46]) and non-sedating (K=4; N=1291; RR 1.60 [1.33, 1.92]) antipsychotic augmentation are clinically important and statistically significant but the effect size is larger for the non-sedating antipsychotics.

Table 103: Study information table for trials included in the meta-analysis of augmenting the antidepressant with another antidepressant or a non-antidepressant agent versus continuing with the antidepressant-only (part 1)

	TeCA + SSRI versus SSRI-only	Lithium + SSRI versus SSRI-only
Total no. of studies (N randomised)	2 (399)	1 (25)
Study ID	Ferreri 2001[1] Licht 2002[2]	Baumann 1996
Country	France[1] Denmark and Iceland[2]	Switzerland
Diagnostic status	DSM-III-R MDD[1] DSM-IV MDD, without psychotic symptoms[2]	DSM-III MDD (single or recurrent), 88%; bipolar disorder; anxiety disorder NOS
Age range (mean)	Range NR (46.6)[1] Range NR (40.3)[2]	Range NR (41.8)
Sex (% female)	74[1] 62[2]	71
Ethnicity (% BME)	NR	NR
Mean age (SD) at first onset of depression	NR[1] 33 (12) [2]	36.2 (13.1)
Mean months (SD) since onset of current episode	7.3 (8.4)[1] Median: 4[2]	4.1 (5.3)
No. (SD) of previous depressive episodes	2.4 (2.2)[1] Median: 2[2]	1.6 (5.7)
Details of inadequate response/treatment resistance	Inadequate response to previous fluoxetine (20mg/day) treatment after ≥6 weeks[1] Inadequate response to 6 weeks of open-label treatment with sertraline (50-100mg/day)[2]	Inadequate response (improvement<50% on HAM-D) to 4-week prospective treatment phase with citalopram (20-60mg/day)
Augmented/previous treatment	Augmented antidepressant: Fluoxetine (20mg/day)[1] Augmented antidepressant: Sertraline (100mg/day) [2]	Augmented antidepressant: Citalopram (40-60mg/day; final mean dose 54mg/day [SD=9])
Baseline severity	HAMD 27.2 (More severe)[1] NR[2]	NR
Intervention details (mean dose)	Mianserin (60mg/day, + fluoxetine 20mg/day)[1] Mianserin (10-30mg/day, + sertraline 100mg/day)[2]	Lithium (800mg/day, target plasma levels 0.5-0.8 mmol/L) + citalopram (40-60mg/day; final mean dose 54mg/day [SD=9])
Comparator details (mean dose, if applicable)	Fluoxetine (20mg/day)[1] Sertraline (100mg/day; + placebo)[2]	Citalopram (40-60mg/day; final mean dose 54mg/day [SD=9])
Treatment length (weeks)	6[1] 5[2]	1

Notes:
Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, NOS=not otherwise specified
[1]Ferreri 2001; [2]Licht 2002

Update 2017

	TeCA + SSRI versus SSRI-only	Lithium + SSRI versus SSRI-only
Note that Ferreri 2001[1] and Licht 2002[2] are three-armed trials and demographics reported here are for all three arms combined		

Table 104: Study information table for trials included in the meta-analysis of augmenting the antidepressant with another antidepressant or a non-antidepressant agent versus continuing with the antidepressant-only (part 2)

	Antipsychotic + SSRI versus SSRI-only	Anticonvulsant + SSRI versus SSRI-only
Total no. of studies (N randomised)	3 (1044)	1 (375)
Study ID	Dunner 2007[1] Fang 2010/2011[2] Thase 2007[3]	Fang 2010/2011
Country	US[1] China[2] US and Canada[3]	China
Diagnostic status	DSM-IV MDD[1,2] DSM-IV MDD (recurrent), without psychotic features, confirmed by the SCID-I[3]	DSM-IV MDD
Age range (mean)	Range NR (44.0) [1] NR[2] 18-65 (44.4) [3]	NR
Sex (% female)	52[1] NR[2] 63[3]	NR
Ethnicity (% BME)	11[1] NR[2] 14[3]	NR
Mean age (SD) at first onset of depression	NR	NR
Mean months (SD) since onset of current episode	NR[1,2] 57.7 (80.9)[3]	NR
No. (SD) of previous depressive episodes	NR	NR
Details of inadequate response/treatment resistance	TRD: Failure to respond to at least 1 previous course of treatment of at least 4 weeks' duration with a clinically appropriate dose of an SSRI or non-SSRI antidepressant (based on self-report), and failure to respond (<30% improvement in MADRS score and continued to have a CGI-S score ≥4 and meet DSM-IV criteria for MDD) to an initial 6-week open-label prospective treatment phase with sertraline[1] TRD: Inadequate response to 2 or more adequate treatments from different classes of	TRD: Inadequate response to 2 or more adequate treatments from different classes of antidepressants in the current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment

Update 2017

	Antipsychotic + SSRI versus SSRI-only	Anticonvulsant + SSRI versus SSRI-only
	antidepressants in the current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment[2] TRD: Documented history of failure to achieve a satisfactory response (based on investigator's clinical judgement) to an antidepressant (except fluoxetine) after at least 6 weeks of therapy at a therapeutic dose (e.g. paroxetine 40mg/day, venlafaxine 150mg/day, bupropion 300mg/day, trazodone 450mg/day), and failure to respond (<25% decrease in HAMD) to an 8-week, open-label prospective fluoxetine (25-50mg/day) therapy lead-in[3]	
Augmented/previous treatment	Augmented antidepressant: Sertraline[1] Paroxetine[2] Fluoxetine[3]	Augmented antidepressant: Paroxetine
Baseline severity	MADRS 29.95 (More severe)[1] NR[2] MADRS 30 (More severe)[3]	NR
Intervention details (mean dose)	Ziprasidone 80mg/day or 160mg/day (combined two fixed dosage arms; mean daily doses 78mg [SD=2.3] and 129.9mg [SD=33.7]) + sertraline 100-200mg/day (mean daily dose 184.3mg [SD=29.7])[1] Risperidone 2mg/day + paroxetine 20mg/day[2] Olanzapine: 6, 12 or 18mg/day (mean modal dose 8.6mg/day) + fluoxetine 50mg/day (mean modal dose 48.8mg/day)[3]	Risperidone 2mg/day + paroxetine 20mg/day
Comparator details (mean dose, if applicable)	Sertraline 100-200mg/day[1] Paroxetine 20mg/day[2] Fluoxetine 50mg/day (mean modal dose 49.5mg/day)[3]	Paroxetine 20mg/day
Treatment length (weeks)	8	8
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression [1]Dunner 2007; [2]Fang 2010/2011; [3]Thase 2007		

Update 2017

	Antipsychotic + SSRI versus SSRI-only	Anticonvulsant + SSRI versus SSRI-only
Note that Dunner 2007[1] and Thase 2007[3] are three-armed trials and demographics reported here are for all three arms combined, and Fang 2010/2011[2] is an eight-armed trial and demographics reported here are for all eight arms combined		

Table 105: Study information table for trials included in the meta-analysis of augmenting the antidepressant with another antidepressant or a non-antidepressant agent versus continuing with the antidepressant-only (part 3)

	Anxiolytic + SSRI versus SSRI-only	SARI + SSRI versus SSRI-only	Thyroid hormone + SSRI versus SSRI-only
Total no. of studies (N randomised)	1 (375)		
Study ID	Fang 2010/2011		
Country	China		
Diagnostic status	DSM-IV MDD		
Age range (mean)	NR		
Sex (% female)	NR		
Ethnicity (% BME)	NR		
Mean age (SD) at first onset of depression	NR		
Mean months (SD) since onset of current episode	NR		
No. (SD) of previous depressive episodes	NR		
Details of inadequate response/treatment resistance	TRD: Inadequate response to 2 or more adequate treatments from different classes of antidepressants in the current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment		
Augmented/previous treatment	Augmented antidepressant: Paroxetine		
Baseline severity	NR		
Intervention details (mean dose)	Buspirone 30mg/day + paroxetine 20mg/day	Trazodone 100mg/day + paroxetine 20mg/day	Thyroid hormone 80mg/day + paroxetine 20mg/day
Comparator details (mean dose, if applicable)	Paroxetine 20mg/day		
Treatment length (weeks)	8		
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression Note that Fang 2010/2011 is an eight-armed trial and demographics reported here are for all eight arms combined			

Update 2017

Table 106: Summary of findings table for augmenting the antidepressant with another antidepressant or a non-antidepressant agent versus continuing with the antidepressant-only

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 Augmenting the antidepressant with another antidepressant/non-antidepressant agent versus continuing with the antidepressant-only				
Remission (TeCA [mianserin] + SSRI versus SSRI-only) Number of people scoring ≤7/8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 5-6 weeks	**Study population** **324 per 1000**	**492 per 1000** (249 to 974)	**RR 1.52** (0.77 to 3.01)	266 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	Moderate **281 per 1000**	**427 per 1000** (216 to 846)				
Remission (antipsychotic + SSRI versus SSRI-only) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D)/≤10 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 8 weeks	**Study population** **209 per 1000**	**234 per 1000** (96 to 575)	**RR 1.12** (0.46 to 2.75)	551 (3 studies)	⊕⊖⊖⊖ **very low**[2,4,5,6]	
	Moderate **168 per 1000**	**188 per 1000** (77 to 462)				
Remission (anticonvulsant + SSRI versus SSRI-only) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population** **467 per 1000**	**485 per 1000** (313 to 761)	**RR 1.04** (0.67 to 1.63)	84 (1 study)	⊕⊖⊖⊖ **very low**[4,6,7]	
	Moderate **467 per 1000**	**486 per 1000** (313 to 761)				
	Study population					

Update 2017

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	Augmenting the antidepressant with another antidepressant/non-antidepressant agent versus continuing with the antidepressant-only				
Remission (anxiolytic + SSRI versus SSRI-only) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**467 per 1000**	**327 per 1000** (196 to 551)	**RR 0.7** (0.42 to 1.18)	91 (1 study)	⊕⊖⊖⊖ **very low**[4,7,8]	
	Moderate					
	467 per 1000	**327 per 1000** (196 to 551)				
Remission (SARI + SSRI versus SSRI-only) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.91** (0.58 to 1.44)	92 (1 study)	⊕⊖⊖⊖ **very low**[4,6,7]	
	467 per 1000	**425 per 1000** (271 to 672)				
	Moderate					
	467 per 1000	**425 per 1000** (271 to 672)				
Remission (thyroid hormone + SSRI versus SSRI-only) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.41** (0.77 to 2.58)	93 (1 study)	⊕⊖⊖⊖ **very low**[3,4,7]	
	267 per 1000	**376 per 1000** (205 to 688)				
	Moderate					
	267 per 1000	**376 per 1000** (206 to 689)				
Response (TeCA [mianserin] + SSRI versus SSRI-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 5-6 weeks	**Study population**		**RR 1.22** (0.69 to 2.15)	266 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,4,6]	
	610 per 1000	**745 per 1000** (421 to 1000)				
	Moderate					
	536 per 1000	**654 per 1000** (370 to 1000)				

Update 2017

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	Augmenting the antidepressant with another antidepressant/non-antidepressant agent versus continuing with the antidepressant-only				
Response (lithium + SSRI versus SSRI-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 1 weeks	**Study population**		**RR 4.2** (1.06 to 16.68)	24 (1 study)	⊕⊖⊖⊖ **very low**[4,9,10]	
	143 per 1000	**600 per 1000** (151 to 1000)				
	Moderate					
	143 per 1000	**601 per 1000** (152 to 1000)				
Response (antipsychotic + SSRI versus SSRI-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 8 weeks	**Study population**		**RR 1.12** (0.61 to 2.07)	551 (3 studies)	⊕⊖⊖⊖ **very low**[2,4,5,6]	
	343 per 1000	**384 per 1000** (209 to 711)				
	Moderate					
	296 per 1000	**332 per 1000** (181 to 613)				
Response (anticonvulsant + SSRI versus SSRI-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.92** (0.67 to 1.27)	84 (1 study)	⊕⊖⊖⊖ **very low**[4,6,7]	
	667 per 1000	**613 per 1000** (447 to 847)				
	Moderate					
	667 per 1000	**614 per 1000** (447 to 847)				
	Study population					

Update 2017

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	Augmenting the antidepressant with another antidepressant/non-antidepressant agent versus continuing with the antidepressant-only				
Response (anxiolytic + SSRI versus SSRI-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**667 per 1000**	**567 per 1000** (407 to 787)	**RR 0.85** (0.61 to 1.18)	91 (1 study)	⊕⊖⊖⊖ **very low**[4,7,8]	
	Moderate					
	667 per 1000	**567 per 1000** (407 to 787)				
Response (SARI + SSRI versus SSRI-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.93** (0.68 to 1.26)	92 (1 study)	⊕⊖⊖⊖ **very low**[4,6,7]	
	667 per 1000	**620 per 1000** (453 to 840)				
	Moderate					
	667 per 1000	**620 per 1000** (454 to 840)				
Response (thyroid hormone + SSRI versus SSRI-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.25** (0.84 to 1.85)	93 (1 study)	⊕⊖⊖⊖ **very low**[3,4,7]	
	467 per 1000	**583 per 1000** (392 to 863)				
	Moderate					
	467 per 1000	**584 per 1000** (392 to 864)				
Response (TeCA [mianserin] + SSRI versus SSRI-only) Number of people rated as much or very much improved on	**Study population**		**RR 1.17** (0.65 to 2.12)	266 (2 studies)	⊕⊖⊖⊖ **very low**[1,4,6,11]	
	743 per 1000	**869 per 1000** (483 to 1000)				
	Moderate					

Update 2017

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	Augmenting the antidepressant with another antidepressant/non-antidepressant agent versus continuing with the antidepressant-only				
Clinical Global Impressions scale (CGI-I) Follow-up: 5-6 weeks	**652 per 1000**	**763 per 1000** (424 to 1000)				
Depression symptomatology (any augmentation agent) Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 6-8 weeks		The mean depression symptomatology (any augmentation agent) in the intervention groups was **0.37 standard deviations lower** (0.55 to 0.2 lower)		531 (3 studies)	⊕⊕⊖⊖ **low**[4,12]	SMD -0.37 (-0.55 to -0.2)
Depression symptomatology (TeCA [mianserin]) + SSRI versus SSRI-only) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 6 weeks		The mean depression symptomatology (TeCA [mianserin]) + SSRI versus SSRI-only) in the intervention groups was **0.66 standard deviations lower** (1.14 to 0.17 lower)		70 (1 study)	⊕⊖⊖⊖ **very low**[4,13,14]	SMD -0.66 (-1.14 to -0.17)
Depression symptomatology (antipsychotic + SSRI versus SSRI-only) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: mean 8 weeks		The mean depression symptomatology (antipsychotic + SSRI versus SSRI-only) in the intervention groups was **0.33 standard deviations lower** (0.52 to 0.15 lower)		461 (2 studies)	⊕⊕⊖⊖ **low**[4,15]	SMD -0.33 (-0.52 to -0.15)
	Study population					

Update 2017

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	Augmenting the antidepressant with another antidepressant/non-antidepressant agent versus continuing with the antidepressant-only				
Discontinuation for any reason (any augmentation agent) Number of participants discontinuing for any reason (including adverse events) Follow-up: 5-8 weeks	**171 per 1000**	**244 per 1000** (183 to 326)	**RR 1.43** (1.07 to 1.91)	734 (4 studies)	⊕⊖⊖⊖ **very low**[4,10,16]	
	Moderate					
	189 per 1000	**270 per 1000** (202 to 361)				
Discontinuation for any reason (TeCA [mianserin] + SSRI versus SSRI-only) Number of participants discontinuing for any reason (including adverse events) Follow-up: 5-6 weeks	**Study population**		**RR 1.43** (0.79 to 2.56)	267 (2 studies)	⊕⊖⊖⊖ **very low**[4,17,18]	
	124 per 1000	**177 per 1000** (98 to 318)				
	Moderate					
	143 per 1000	**204 per 1000** (113 to 366)				
Discontinuation for any reason (antipsychotic + SSRI versus SSRI-only) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 8 weeks	**Study population**		**RR 1.44** (1.03 to 2)	467 (2 studies)	⊕⊖⊖⊖ **very low**[4,10,19]	
	199 per 1000	**287 per 1000** (205 to 398)				
	Moderate					
	222 per 1000	**320 per 1000** (229 to 444)				
Discontinuation due to adverse events (any augmentation agent) Number of participants	**Study population**		**RR 6.19** (2.65 to 14.47)	537 (3 studies)	⊕⊖⊖⊖ **very low**[4,10,12]	
	19 per 1000	**117 per 1000** (50 to 274)				
	Moderate					

Update 2017

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	Augmenting the antidepressant with another antidepressant/non-antidepressant agent versus continuing with the antidepressant-only				
discontinuing due to adverse events Follow-up: 6-8 weeks	**0 per 1000**	**0 per 1000** (0 to 0)				
Discontinuation due to adverse events (TeCA [mianserin] + SSRI versus SSRI-only) Number of participants discontinuing due to adverse events Follow-up: mean 6 weeks	**Study population**		**RR 5.91** (0.29 to 118.78)	70 (1 study)	⊕⊖⊖⊖ **very low**[4,20,21]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				
Discontinuation due to adverse events (antipsychotic + SSRI versus SSRI-only) Number of participants discontinuing due to adverse events Follow-up: mean 8 weeks	**Study population**		**RR 6.22** (2.57 to 15.07)	467 (2 studies)	⊕⊖⊖⊖ **very low**[4,10,19]	
	22 per 1000	**138 per 1000** (57 to 333)				
	Moderate					
	12 per 1000	**75 per 1000** (31 to 181)				

Notes:
[1] Unclear blinding of intervention administration, and unclear blinding or non-blind outcome assessment
[2] I-squared>50%
[3] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[4] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical company
[5] Unclear or high risk of bias associated with randomisation method and unclear method of allocation concealment, unclear blinding of intervention administration and outcome assessment, and unclear risk of attrition bias (drop-out>20% and some differences between groups but ITT analysis used) in studies contributing>50% to weighting in analysis
[6] 95% CI crosses line of no effect and both threshold for clinically important harm (RR 0.75) and threshold for clinically important benefit (RR 1.25)
[7] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and unclear blinding of intervention administration and outcome assessment
[8] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[9] Unclear randomisation method and method of allocation concealment, and unclear blinding of

Update 2017

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	Augmenting the antidepressant with another antidepressant/non-antidepressant agent versus continuing with the antidepressant-only				
intervention administration and outcome assessment [10] Events<300 [11] I-squared>80% [12] Unclear randomisation method and method of allocation concealment, unclear or non-blind intervention administration and outcome assessment, and unclear risk of attrition bias (drop-out>20% and some differences between groups but ITT analysis used), in studies contributing >50% to weighting in analysis [13] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administrator. Outcome assessment was non-blind. There was also an unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used) [14] N<400 [15] Unclear randomisation method and method of allocation concealment, and unclear or non-blind intervention administration. There was also an unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used) in studies contributing >50% to weighing in analysis [16] Unclear randomisation method and method of allocation concealment, and unclear or non-blind intervention administration, in studies contributing >50% to weighting in analysis [17] Unclear blinding of intervention administration [18] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25) [19] Unclear randomisation method and method of allocation concealment, and unclear or non-blind intervention administration [20] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administrator [21] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)						

Update 2017

Table 107: Study information table for trials included in the meta-analysis of augmenting the antidepressant with lithium versus 'other' augmentation agents (head-to-head comparisons)

	Lithium versus TCA	Lithium versus antipsychotic	Lithium versus thyroid hormone	Lithium versus anticonvulsant
Total no. of studies (N randomised)	2 (142)	2 (708)	3 (229)	1 (34)
Study ID	Fava 1994a[1] Fava 2002[2]	Bauer 2010/2013[3] Doree 2007[4]	Joffe 1993[5] Joffe 2006[6] Nierenberg 2006[7]	Schindler 2007
Country	US	Australia, Austria, Belgium, Bulgaria, Germany, Hungary, Italy,	Canada[5,6] US[7]	Germany

	Lithium versus TCA	Lithium versus antipsychotic	Lithium versus thyroid hormone	Lithium versus anticonvulsant
		Portugal, Romania, Slovakia, Spain and the UK[3] Canada[4]		
Diagnostic status	DSM-III-R MDD	DSM-IV diagnosis of MDD (single or recurrent episode), confirmed by the Mini International Neuropsychiatric Interview (MINI)[3] DSM‑IV‑TR unipolar MDD, without psychotic features[4]	RDC criteria for unipolar, nonpsychotic MDD[5] DSM-IV criteria for nonpsychotic, unipolar MDD[6] DSM-IV nonpsychotic MDD[7]	DSM-IV-TR non-psychotic, unipolar major depressive episode
Age range (mean)	18-65 (39.6)[1] Range NR (41.6)[2]	NR[3] Range NR (50.8)[4]	Range NR (37.4)[5] 23-52 (39.2)[6] Range NR (42.0)[7]	Range NR (47.7)
Sex (% female)	61[1] 49[2]	NR[3] 60[4]	61[5] 83[6] 58[7]	50
Ethnicity (% BME)	NR	NR	NR[5,6] 17[7]	NR
Mean age (SD) at first onset of depression	NR	NR	NR[5,6] 23.5 (13.7)[7]	NR
Mean months (SD) since onset of current episode	NR	6 (3.8)[3] NR[4]	NR[5,6] 29.5 (74.2)[7]	7.4 (2.6)
No. (SD) of previous depressive episodes	NR	4.0 (6.0)[3] NR[4]	NR[5,6] 7.4 (14.6)[7]	2.9 (1.2)
Details of inadequate response/treatment resistance	Inadequate response (defined as failure to achieve a 50% or greater reduction in HAMD score and a HAMD score of ≥10) to 8 weeks of open-label treatment with fluoxetine (20mg/day)	Patients were required to have Stage I or II TRD, 50% of participants fell into each category (defined as: Stage I-failure of ≥1 adequate trial of one major class of AD [citalopram, escitalopram, paroxetine, sertraline or venlafaxine]; Stage II-failure of	Inadequate response (had a score≥16 on the 17-item HAMD) to a previous adequate trial of desipramine hydrochloride (90%) or imipramine hydrochloride (10%) ≥5 weeks (at a minimum dose of 2.5mg/kg of body weight per day)[5]	TRD: Inadequate response (<50%-reduction of initial HRSD) to at least two trials of different classes of antidepressants for a duration of at least 6 weeks

Update 2017

	Lithium versus TCA	Lithium versus antipsychotic	Lithium versus thyroid hormone	Lithium versus anticonvulsant
		adequate trials of two different classes of AD, the most recent of which must have been an AD listed for patients with Stage I TRD). An inadequate response was defined as not achieving remission from depressive symptoms after receiving at least a minimum effective dose of an AD with ≥1 dose increase for ≥28 days prior to the study [3] Inadequate response after 4 weeks of treatment with an antidepressant at the maximal recommended dose[4]	Inadequate response to a trial of antidepressants at usual dosages (moclobemide 600 to 750 mg daily[8], nefazodone 150 to 300 mg daily, paroxetine 20 to 60 mg daily, sertraline 100 to 200 mg daily, fluoxetine 30 to 40 mg daily, fluvoxamine 150 to 300 mg daily, and venlafaxine 187.5 to 375 mg daily) for at least 5 weeks[6] TRD: Inadequate response (not achieved remission or who were intolerant) to an initial prospective treatment with citalopram and a second switch or augmentation trial[7]	
Augmented/previous treatment	Augmented antidepressant: Fluoxetine (20mg/day)	Augmented antidepressant: 66% SSRI; 36% venlafaxine; 8% other AD[3] Augmented antidepressant (55% receiving two antidepressants): 40% mirtazapine (30-45mg); 25% venlafaxine (187.5-300mg); 20% paroxetine (20-50mg); 20% trazodone (25-200mg); 15% citalopram (40-60mg); 15% buproprion (400-600mg); 10% sertraline	Augmented antidepressant: 90% desipramine; 10% imipramine (mean dose of desipramine or imipramine 201mg/day [range 150-300mg/day]) [5] Augmented antidepressant: 78% SSRI[6] Augmented antidepressant: citalopram and bupropion (24%; mean dose 326.5 mg); bupropion (21%; mean dose 395.0 mg);	Type of augmented antidepressant NR (the prior antidepressive medication was continued throughout the study, prior augmentation strategies were discontinued)

Update 2017

	Lithium versus TCA	Lithium versus antipsychotic	Lithium versus thyroid hormone	Lithium versus anticonvulsant
		(200mg); 5% nefazadone (300mg)[4]	venlafaxine (21%; mean dose 316.3 mg); citalopram and buspirone (19%; mean dose 46.7 mg); sertraline (15%; mean dose 183.3 mg)[7]	
Baseline severity	HAMD 14.5 (Less severe)[1] HAMD 16.6 (Less severe)[2]	MADRS 33.3 (More severe)[3] NR[4]	HAMD 19.5 (Less severe)[5,6] HAMD 18.1 (Less severe)[7]	HAMD 22.1 (Less severe)
Intervention details (mean dose)	Lithium 300-600mg/day	Lithium 450-900mg/day (target plasma level: 0.6–1.2mmol/L; mean dose 882mg/day [SD=212])[3] Lithium 600mg/day, target plasma levels 0.8–1.2 mmol/L (mean final plasma level 0.66 mmol/L)[4]	Lithium 900-1200mg/day (target plasma level 0.55 nmol/L)[5] Lithium 600-900mg/day[6] Lithium 225-900mg/day (mean final dose 859.8mg/day)[7]	Lithium target plasma level 0.6–0.8mmol/l (mean final plasma level 0.71mmol/l)
Comparator details (mean dose, if applicable)	Desipramine 25-50mg/day	Quetiapine extended-release (XR) 200-300mg/day (titrated upwards from 50mg/day to 300mg/day in first week and titrated downwards if necessary; mean dose 242mg/day [SD=54])[3] Quetiapine 400-800mg/day (titrated up to 400mg within the first week; mean final dose 430mg [range 300-700mg])[4]	Liothyronine sodium (triiodothyronine, T3) 37.5µg/day[5,6] Thyroid hormone (T3) 25-50 µg/day (mean final dose 45.2µg/day)[7]	Lamotrigine 25-250mg/day (mean final dose 152.94 mg/day)
Treatment length (weeks)	4	6[3] 8[4]	2[5,6] 14[7]	8
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression				

Update 2017

	Lithium versus TCA	Lithium versus antipsychotic	Lithium versus thyroid hormone	Lithium versus anticonvulsant

[1]Fava 1994a; [2]Fava 2002; [3]Bauer 2010/2013; [4]Doree 2007; [5]Joffe 1993; [6]Joffe 2006; [7]Nierenberg 2006

Note that Bauer 2010/2013[3], Fava 1994a[1], Fava 2002[2] and Joffe 1993[5] are three-armed trials and demographics reported here are for all three arms combined, and Joffe 2006[6] is a four-armed trial and demographics reported here for all four arms combined

[8]Note that the previous inadequate response was to a higher than licensed dose range for moclobemide (300-600mg/day) and for some drugs in the table the dose ranges used in the studies were greater than the licensed dose ranges in the Summaries of Product Characteristics (SPCs).

Table 108: Summary of findings table for augmenting the antidepressant with lithium versus 'other' augmentation agents (head-to-head comparisons)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	'other' augmentation agent	Augmenting the antidepressant with lithium				
Remission (lithium versus any 'other' augmentation agent) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D)/≤8/10 on Montgomery Asberg Depression Rating Scale (MADRS)/<7 on HAM-D AND responding (≥50% improvement on HAM-D) Follow-up: 2-14 weeks	**Study population**		**RR 0.79** (0.63 to 0.99)	774 (7 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	314 per 1000	**248 per 1000** (198 to 311)				
	Moderate					
	294 per 1000	**232 per 1000** (185 to 291)				
Remission (lithium versus TCA) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 4 weeks	**Study population**		**RR 0.88** (0.45 to 1.74)	94 (2 studies)	⊕⊖⊖⊖ **very low**[3,4,5]	
	283 per 1000	**249 per 1000** (127 to 492)				
	Moderate					
	272 per 1000	**239 per 1000** (122 to 473)				
Remission (lithium versus antipsychotic) Number of people	**Study population**		**RR 0.65** (0.31 to 1.39)	470 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,5,6]	
	339 per 1000	**220 per 1000** (105 to 471)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk 'other' augmentation agent	Corresponding risk Augmenting the antidepressant with lithium				
scoring ≤8/<10 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 6-8 weeks	**Moderate** **559 per 1000**	**363 per 1000** (173 to 777)				
Remission (lithium versus thyroid hormone [T3]) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D)/<7 AND responding (≥50% improvement on HAM-D) Follow-up: 2-14 weeks	**Study population** **278 per 1000** **Moderate** **329 per 1000**	**200 per 1000** (117 to 339) **237 per 1000** (138 to 401)	**RR 0.72** (0.42 to 1.22)	176 (2 studies)	⊕⊖⊖⊖ **very low**[3,7,8]	
Remission (lithium versus anticonvulsant [lamotrigine]) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population** **235 per 1000** **Moderate** **235 per 1000**	**176 per 1000** (47 to 673) **176 per 1000** (47 to 672)	**RR 0.75** (0.2 to 2.86)	34 (1 study)	⊕⊖⊖⊖ **very low**[4,5]	
Response (lithium versus any 'other' augmentation agent) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Montgomery Asberg Depression Rating Scale (MADRS)/Quick Inventory of Depressive Symptomatology (QIDS) Follow-up: 6-14 weeks	**Study population** **468 per 1000** **Moderate** **527 per 1000**	**426 per 1000** (365 to 506) **480 per 1000** (411 to 569)	**RR 0.91** (0.78 to 1.08)	646 (4 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	'other' augmentation agent	Augmenting the antidepressant with lithium				
Response (lithium versus antipsychotic) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 6-8 weeks	**Study population**		**RR 0.89** (0.63 to 1.25)	470 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,8]	
	536 per 1000	**477 per 1000** (337 to 669)				
	Moderate					
	662 per 1000	**589 per 1000** (417 to 827)				
Response (lithium versus thyroid hormone [T3]) Number of people showing ≥50% improvement on Quick Inventory of Depressive Symptomatology (QIDS) Follow-up: mean 14 weeks	**Study population**		**RR 0.68** (0.35 to 1.36)	142 (1 study)	⊕⊖⊖⊖ **very low**[3,5,9]	
	233 per 1000	**158 per 1000** (82 to 317)				
	Moderate					
	233 per 1000	**158 per 1000** (82 to 317)				
Response (lithium versus anticonvulsant [lamotrigine]) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.78** (0.38 to 1.6)	34 (1 study)	⊕⊖⊖⊖ **very low**[4,5]	
	529 per 1000	**413 per 1000** (201 to 847)				
	Moderate					
	529 per 1000	**413 per 1000** (201 to 846)				
Response (lithium versus antipsychotic) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 6 weeks	**Study population**		**RR 0.9** (0.78 to 1.04)	450 (1 study)	⊕⊖⊖⊖ **very low**[2,3,10]	
	668 per 1000	**601 per 1000** (521 to 695)				
	Moderate					
	668 per 1000	**601 per 1000** (521 to 695)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	'other' augmentation agent	Augmenting the antidepressant with lithium				
Depression symptomatology (lithium versus any 'other' augmentation agent) Hamilton Rating Scale for Depression (HAM-D; change score)/Quick Inventory of Depressive Symptomatology (QIDS; change score) Follow-up: 2-14 weeks		The mean depression symptomatology (lithium versus any 'other' augmentation agent) in the intervention groups was **0.14 standard deviations higher** (0.14 lower to 0.42 higher)		304 (5 studies)	⊕⊖⊖⊖ **very low**[3,11,12]	SMD 0.14 (-0.14 to 0.42)
Depression symptomatology (lithium versus TCA) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 4 weeks		The mean depression symptomatology (lithium versus TCA) in the intervention groups was **0.09 standard deviations lower** (0.49 lower to 0.32 higher)		94 (2 studies)	⊕⊖⊖⊖ **very low**[3,4,12]	SMD -0.09 (-0.49 to 0.32)
Depression symptomatology (lithium versus thyroid hormone [T3]) Hamilton Rating Scale for Depression (HAM-D; change score)/Quick Inventory of Depressive Symptomatology (QIDS; change score) Follow-up: 2-14 weeks		The mean depression symptomatology (lithium versus thyroid hormone [t3]) in the intervention groups was **0.15 standard deviations higher** (0.14 lower to 0.45 higher)		176 (2 studies)	⊕⊖⊖⊖ **very low**[3,11,12]	SMD 0.15 (-0.14 to 0.45)
Depression symptomatology (lithium versus anticonvulsant [lamotrigine])		The mean depression symptomatology (lithium versus anticonvulsant		34 (1 study)	⊕⊕⊖⊖ **low**[4,13]	SMD 0.81 (0.11 to 1.51)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	'other' augmentation agent	Augmenting the antidepressant with lithium				
Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 8 weeks		[lamotrigine]) in the intervention groups was **0.81 standard deviations higher** (0.11 to 1.51 higher)				
Discontinuation for any reason (lithium versus any 'other' augmentation agent) Number of participants discontinuing for any reason (including adverse events) Follow-up: 2-8 weeks	**Study population**		**RR 1.29** (0.91 to 1.84)	662 (7 studies)	⊕⊖⊖⊖ **very low**[3,10,14]	
	136 per 1000	**175 per 1000** (124 to 250)				
	Moderate					
	118 per 1000	**152 per 1000** (107 to 217)				
Discontinuation for any reason (lithium versus TCA) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 4 weeks	**Study population**		**RR 0.83** (0.33 to 2.11)	94 (2 studies)	⊕⊖⊖⊖ **very low**[3,15,16]	
	174 per 1000	**144 per 1000** (57 to 367)				
	Moderate					
	199 per 1000	**165 per 1000** (66 to 420)				
Discontinuation for any reason (lithium versus antipsychotic) Number of participants discontinuing for any reason (including adverse events) Follow-up: 6-8 weeks	**Study population**		**RR 1.66** (0.57 to 4.79)	480 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,16]	
	145 per 1000	**241 per 1000** (83 to 696)				
	Moderate					
	76 per 1000	**126 per 1000** (43 to 364)				
Discontinuation for any reason (lithium versus thyroid hormone [T3]) Number of participants discontinuing for any reason (including adverse events)	**Study population**		**RR 2.84** (0.12 to 65.34)	54 (2 studies)	⊕⊖⊖⊖ **very low**[15,16]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk 'other' augmentation agent	Corresponding risk Augmenting the antidepressant with lithium				
Follow-up: mean 2 weeks						
Discontinuation for any reason (lithium versus anticonvulsant [lamotrigine]) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 8 weeks	**Study population**		**RR 1** (0.16 to 6.3)	34 (1 study)	⊕⊖⊖⊖ **very low**[15,16]	
	118 per 1000	**118 per 1000** (19 to 741)				
	Moderate					
	118 per 1000	**118 per 1000** (19 to 743)				
Discontinuation due to adverse events (lithium versus any 'other' augmentation agent) Number of participants discontinuing due to adverse events Follow-up: 2-14 weeks	**Study population**		**RR 1.27** (0.61 to 2.64)	736 (7 studies)	⊕⊖⊖⊖ **very low**[3,16,17]	
	86 per 1000	**110 per 1000** (53 to 228)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				
Discontinuation due to adverse events (lithium versus TCA) Number of participants discontinuing due to adverse events Follow-up: mean 4 weeks	**Study population**		**RR 0.43** (0.04 to 4.16)	26 (1 study)	⊕⊖⊖⊖ **very low**[3,15,16]	
	167 per 1000	**72 per 1000** (7 to 693)				
	Moderate					
	167 per 1000	**72 per 1000** (7 to 695)				
Discontinuation due to adverse events (lithium versus antipsychotic) Number of participants discontinuing due to adverse events Follow-up: 6-8 weeks	**Study population**		**RR 0.83** (0.46 to 1.48)	480 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,16]	
	95 per 1000	**79 per 1000** (44 to 141)				
	Moderate					
	50 per 1000	**42 per 1000** (23 to 74)				
	Study population					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	'other' augmentation agent	Augmenting the antidepressant with lithium				
Discontinuation due to adverse events (lithium versus thyroid hormone [T3]) Number of participants discontinuing due to adverse events Follow-up: 2-14 weeks	**70 per 1000** **Moderate** **0 per 1000**	**171 per 1000** (77 to 380) **0 per 1000** (0 to 0)	**RR 2.44** (1.1 to 5.43)	196 (3 studies)	⊕⊖⊖⊖ **very low**[2,3,17]	
Discontinuation due to adverse events (lithium versus anticonvulsant [lamotrigine]) Number of participants discontinuing due to adverse events Follow-up: mean 8 weeks	See comment	See comment	Not estimable	34 (1 study)	⊕⊕⊖⊖ **low**[2,15]	

Notes:
[1] Unclear method of allocation concealment and unclear or non-blind intervention administration in studies contributing >50% to weighting in analysis
[2] Events<300
[3] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical company
[4] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration and outcome assessment
[5] 95% CI crosses line of no effect and threshold for clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[6] I-squared>50%
[7] High risk of bias associated with randomisation method due to significant difference between groups at baseline (in studies contributing >50% to weighting in analysis) and unclear method of allocation concealment and unclear blinding of intervention administration
[8] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[9] High risk of bias associated with randomisation method due to significant difference between groups at baseline and unclear method of allocation concealment and unclear blinding of intervention administration
[10] Unclear method of allocation concealment and non-blind intervention administration
[11] Risk associated with randomisation method is high or unclear, the method of allocation concealment is unclear, and blinding of intervention administration and outcome assessment is unclear, in studies contributing to >50% of weighting in analysis
[12] N<400
[13] 95% CI crosses both line of no effect and threshold for clinically important harm (SMD 0.5)
[14] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)
[15] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	'other' augmentation agent	Augmenting the antidepressant with lithium				

[16] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[17] Risk associated with randomisation method is high or unclear, the method of allocation concealment is unclear, and blinding of intervention administration is unclear, in studies contributing to >50% of weighting in analysis

Table 109: Study information table for trials included in the meta-analysis of augmenting the antidepressant with an antipsychotic versus 'other' augmentation agents (head-to-head comparisons)

	Antipsychotic versus anticonvulsant	Antipsychotic versus anxiolytic	Antipsychotic versus thyroid hormone	Antipsychotic versus SARI
Total no. of studies (N randomised)	1 (375)			
Study ID	Fang 2010/2011			
Country	China			
Diagnostic status	DSM-IV MDD			
Age range (mean)	NR			
Sex (% female)	NR			
Ethnicity (% BME)	NR			
Mean age (SD) at first onset of depression	NR			
Mean months (SD) since onset of current episode	NR			
No. (SD) of previous depressive episodes	NR			
Details of inadequate response/treatment resistance	TRD: Inadequate response to 2 or more adequate treatments from different classes of antidepressants in the current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment			
Augmented/previous treatment	Augmented antidepressant: Paroxetine (20mg/day)			
Baseline severity	NR			
Intervention details (mean dose)	Risperidone 2mg/day			
Comparator details (mean dose, if applicable)	Sodium valproate 600mg/day	Buspirone 30mg/day	Thyroid hormone 80mg/day	Trazodone 100mg/day
Treatment length (weeks)	8			
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression				

Update 2017

	Antipsychotic versus anticonvulsant	Antipsychotic versus anxiolytic	Antipsychotic versus thyroid hormone	Antipsychotic versus SARI
Note that Fang2010/2011 is an eight-armed trial and demographics reported here are for all eight arms combined				

Table 110: Summary of findings table for augmenting the antidepressant with an antipsychotic versus 'other' augmentation agents (head-to-head comparisons)

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 'other' augmentation agent	Augmenting the antidepressant with an antipsychotic				
Remission (antipsychotic versus anticonvulsant) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.55** (0.31 to 0.98)	84 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	487 per 1000	**268 per 1000** (151 to 477)				
	Moderate					
	487 per 1000	**268 per 1000** (151 to 477)				
Remission (antipsychotic versus anxiolytic) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.82** (0.43 to 1.55)	91 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	326 per 1000	**267 per 1000** (140 to 505)				
	Moderate					
	326 per 1000	**267 per 1000** (140 to 505)				
Remission (antipsychotic versus thyroid hormone) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.71** (0.39 to 1.3)	93 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	375 per 1000	**266 per 1000** (146 to 487)				
	Moderate					
	375 per 1000	**266 per 1000** (146 to 487)				
Remission (antipsychotic versus SARI) Number of people	**Study population**		**RR 0.63** (0.35 to 1.13)	92 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	
	426 per 1000	**268 per 1000** (149 to 481)				

Update 2017

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 'other' augmentation agent	Augmenting the antidepressant with an antipsychotic				
scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Moderate**					
	267 per 1000	**168 per 1000** (93 to 302)				
Response (antipsychotic versus anticonvulsant) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.76** (0.51 to 1.13)	84 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	
	615 per 1000	**468 per 1000** (314 to 695)				
	Moderate					
	615 per 1000	**467 per 1000** (314 to 695)				
Response (antipsychotic versus anxiolytic) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.83** (0.55 to 1.23)	91 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	
	565 per 1000	**469 per 1000** (311 to 695)				
	Moderate					
	565 per 1000	**469 per 1000** (311 to 695)				
Response (antipsychotic versus thyroid hormone) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.8** (0.54 to 1.19)	93 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	
	583 per 1000	**467 per 1000** (315 to 694)				
	Moderate					
	583 per 1000	**466 per 1000** (315 to 694)				
Response (antipsychotic versus SARI) Number of people	**Study population**		**RR 0.76** (0.51 to 1.11)	92 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	
	617 per 1000	**469 per 1000** (315 to 685)				

Update 2017

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 'other' augmentation agent	**Augmenting the antidepressant with an antipsychotic**				
showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Moderate**					
	467 per 1000	**355 per 1000** (238 to 518)				

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and unclear blinding of intervention administration and outcome assessment
[2] Events<300
[3] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical company
[4] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[5] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)

Table 111: Study information table for trials included in the meta-analysis of augmenting the antidepressant with an anticonvulsant versus 'other' augmentation agents (head-to-head comparisons)

	Anticonvulsant versus anxiolytic	Anticonvulsant versus SARI	Anticonvulsant versus thyroid hormone
Total no. of studies (N randomised)	1 (375)		
Study ID	Fang 2010/2011		
Country	China		
Diagnostic status	DSM-IV MDD		
Age range (mean)	NR		
Sex (% female)	NR		
Ethnicity (% BME)	NR		
Mean age (SD) at first onset of depression	NR		
Mean months (SD) since onset of current episode	NR		
No. (SD) of previous depressive episodes	NR		
Details of inadequate response/treatment resistance	TRD: Inadequate response to 2 or more adequate treatments from different classes of antidepressants in the current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment		
Augmented/previous treatment	Augmented antidepressant: Paroxetine 20mg/day		
Baseline severity	NR		

Update 2017

	Anticonvulsant versus anxiolytic	Anticonvulsant versus SARI	Anticonvulsant versus thyroid hormone
Intervention details (mean dose)	Sodium valproate 600mg/day		
Comparator details (mean dose, if applicable)	Buspirone 30mg/day	Trazodone 100mg/day	Thyroid hormone 80mg/day
Treatment length (weeks)	8		

Notes:
Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression
Note that Fang 2010/2011 is an eight-armed trial and demographics reported here are for all eight arms combined

Table 112: Summary of findings table for augmenting the antidepressant with an anticonvulsant versus 'other' augmentation agents (head-to-head comparisons)

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 'other' augmentation agent	Augmenting the antidepressant with an anticonvulsant				
Remission (anticonvulsant versus anxiolytic) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.49** (0.88 to 2.53)	85 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	326 per 1000	**486 per 1000** (287 to 825)				
	Moderate					
	326 per 1000	**486 per 1000** (287 to 825)				
Remission (anticonvulsant versus SARI) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.14** (0.72 to 1.82)	86 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	426 per 1000	**485 per 1000** (306 to 774)				
	Moderate					
	426 per 1000	**486 per 1000** (307 to 775)				
Remission (anticonvulsant versus thyroid hormone) Number of people scoring ≤7 on	**Study population**		**RR 1.3** (0.8 to 2.11)	87 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	375 per 1000	**488 per 1000** (300 to 791)				
	Moderate					

Update 2017

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 'other' augmentation agent	Augmenting the antidepressant with an anticonvulsant				
Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**375 per 1000**	**488 per 1000** (300 to 791)				
Response (anticonvulsant versus anxiolytic) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.09** (0.76 to 1.55)	85 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	565 per 1000	**616 per 1000** (430 to 876)				
	Moderate					
	565 per 1000	**616 per 1000** (429 to 876)				
Response (anticonvulsant versus SARI) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1** (0.71 to 1.39)	86 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	617 per 1000	**617 per 1000** (438 to 858)				
	Moderate					
	617 per 1000	**617 per 1000** (438 to 858)				
Response (anticonvulsant versus thyroid hormone) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.05** (0.75 to 1.49)	87 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	583 per 1000	**612 per 1000** (437 to 869)				
	Moderate					
	583 per 1000	**612 per 1000** (437 to 869)				

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and unclear blinding of intervention administration and outcome assessment
[2] 95% crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical company
[4] 95% CI crosses line of no effect and both threshold for clinically important harm (RR 0.75) and for clinically important benefit (RR 1.25)

Update 2017

Table 113: Study information table for trials included in the meta-analysis of augmenting the antidepressant with an anxiolytic versus 'other' augmentation agents (head-to-head comparisons)

	Anxiolytic versus atypical antidepressant	Anxiolytic versus SARI	Anxiolytic versus thyroid hormone
Total no. of studies (N randomised)	1 (565)	1 (375)	
Study ID	Trivedi 2006	Fang 2010/2011	
Country	US	China	
Diagnostic status	DSM-IV nonpsychotic MDD	DSM-IV MDD	
Age range (mean)	Range NR (41.1)	NR	
Sex (% female)	59	NR	
Ethnicity (% BME)	22	NR	
Mean age (SD) at first onset of depression	25.2 (14.0)	NR	
Mean months (SD) since onset of current episode	27.1 (55.6)	NR	
No. (SD) of previous depressive episodes	6.5 (13.3)	NR	
Details of inadequate response/treatment resistance	Inadequate response (without remission [HAMD>7]) to a mean of 11.9 weeks of citalopram therapy (mean final dose 55mg/day)	TRD: Inadequate response to 2 or more adequate treatments from different classes of antidepressants in the current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment	
Augmented/previous treatment	Augmented antidepressant: Citalopram (mean dose 55mg/day)	Augmented antidepressant: Paroxetine 20mg/day	
Baseline severity	HAMD 15.8 (Less severe)	NR	
Intervention details (mean dose)	Buspirone 15-60mg/day (mean final dose 40.9 mg/day)	Buspirone 30mg/day	
Comparator details (mean dose, if applicable)	Bupropion Sustained Release 200-400mg/day (mean final dose 267.5 mg/day)	Trazodone 100mg/day	Thyroid hormone 80mg/day
Treatment length (weeks)	6	8	
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression Note that Fang 2010/2011 is an eight-armed trial and demographics reported here are for all eight arms combined			

Update 2017

Table 114: Summary of findings table for augmenting the antidepressant with an anxiolytic versus 'other' augmentation agents (head-to-head comparisons)

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 'other' augmentation agent	Augmenting the antidepressant with an anxiolytic				
Remission (anxiolytic versus atypical antidepressant) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 6 weeks	**Study population**		**RR 1.01** (0.79 to 1.3)	565 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	297 per 1000	**300 per 1000** (235 to 387)				
	Moderate					
	298 per 1000	**301 per 1000** (235 to 387)				
Remission (anxiolytic versus SARI) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.77** (0.45 to 1.3)	93 (1 study)	⊕⊖⊖⊖ **very low**[3,4,5]	
	426 per 1000	**328 per 1000** (191 to 553)				
	Moderate					
	426 per 1000	**328 per 1000** (192 to 554)				
Remission (anxiolytic versus thyroid hormone) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.87** (0.5 to 1.51)	94 (1 study)	⊕⊖⊖⊖ **very low**[3,4,5]	
	375 per 1000	**326 per 1000** (188 to 566)				
	Moderate					
	375 per 1000	**326 per 1000** (188 to 566)				
Response (anxiolytic versus atypical antidepressant) Number of people showing ≥50% improvement on Quick Inventory of Depressive Symptomatology (QIDS)	**Study population**		**RR 0.85** (0.66 to 1.1)	565 (1 study)	⊕⊖⊖⊖ **very low**[1,3,6]	
	315 per 1000	**268 per 1000** (208 to 347)				
	Moderate					
	315 per 1000	**268 per 1000** (208 to 347)				

Update 2017

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 'other' augmentation agent	Augmenting the antidepressant with an anxiolytic				
Follow-up: mean 6 weeks						
Response (anxiolytic versus SARI) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.92** (0.65 to 1.29)	93 (1 study)	⊕⊖⊖⊖ **very low**[3,4,5]	
	617 per 1000	**568 per 1000** (401 to 796)				
	Moderate					
	617 per 1000	**568 per 1000** (401 to 796)				
Response (anxiolytic versus thyroid hormone) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.97** (0.68 to 1.37)	94 (1 study)	⊕⊖⊖⊖ **very low**[3,4,5]	
	583 per 1000	**566 per 1000** (397 to 799)				
	Moderate					
	583 per 1000	**566 per 1000** (396 to 799)				
Depression symptomatology (anxiolytic versus atypical antidepressant) Quick Inventory of Depressive Symptomatology (QIDS; change score) Follow-up: mean 6 weeks		The mean depression symptomatology (anxiolytic versus atypical antidepressant) in the intervention groups was **0.17 standard deviations higher** (0.01 to 0.34 higher)		565 (1 study)	⊕⊖⊖⊖ **very low**[1,3]	SMD 0.17 (0.01 to 0.34)
Discontinuation due to adverse events (anxiolytic versus atypical antidepressant) Number of participants discontinuing due to adverse events	**Study population**		**RR 1.64** (1.12 to 2.41)	565 (1 study)	⊕⊖⊖⊖ **very low**[1,3,7]	
	125 per 1000	**206 per 1000** (141 to 302)				
	Moderate					
	125 per 1000	**205 per 1000** (140 to 301)				

Update 2017

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 'other' augmentation agent	Augmenting the antidepressant with an anxiolytic				
Follow-up: mean 6 weeks						

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline and unclear method of allocation concealment, and unclear blinding of intervention administration
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical company
[4] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and unclear blinding of intervention administration and outcome assessment
[5] 95% CI crosses line of no effect and both threshold for clinically important harm (RR 0.75) and for clinically important benefit (RR 1.25)
[6] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[7] Events<300

Table 115: Study information table for trials included in the meta-analysis of augmenting the antidepressant with a thyroid hormone versus 'other' augmentation agents (head-to-head comparisons)

	Thyroid hormone versus SARI
Total no. of studies (N randomised)	1 (375)
Study ID	Fang 2010/2011
Country	China
Diagnostic status	DSM-IV MDD
Age range (mean)	NR
Sex (% female)	NR
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Details of inadequate response/treatment resistance	TRD: Inadequate response to 2 or more adequate treatments from different classes of antidepressants in the current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment
Augmented/previous treatment	Augmented antidepressant: Paroxetine 20mg/day
Baseline severity	NR
Intervention details (mean dose)	Thyroid hormone 80mg/day
Comparator details (mean dose, if applicable)	Trazodone 100mg/day
Treatment length (weeks)	8
Notes:	

Update 2017

	Thyroid hormone versus SARI
Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression Note that Fang 2010/2011 is an eight-armed study and demographics reported here are for all eight arms combined	

Table 116: Summary of findings table for augmenting the antidepressant with a thyroid hormone versus 'other' augmentation agents (head-to-head comparisons)

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 'other' augmentation agent	Augmenting the antidepressant with a thyroid hormone				
Remission (thyroid hormone versus SARI) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.88** (0.54 to 1.44)	95 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	426 per 1000	**374 per 1000** (230 to 613)				
	Moderate					
	426 per 1000	**375 per 1000** (230 to 613)				
Response (thyroid hormone versus SARI) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.95** (0.68 to 1.31)	95 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	617 per 1000	**586 per 1000** (420 to 808)				
	Moderate					
	617 per 1000	**586 per 1000** (420 to 808)				

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and unclear blinding of intervention administration and outcome assessment
[2] 95% CI crosses line of no effect and both threshold for clinically important harm (RR 0.75) and for clinically important benefit (RR 1.25)
[3] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical company

Update 2017

Table 117: Study information table for trials included in the meta-analysis of augmenting the antidepressant with a psychological intervention versus attention-placebo

	Mindfulness-based cognitive therapy (MBCT) versus attention-placebo
Total no. of studies (N randomised)	2 (223)
Study ID	Chiesa 2015[1] Eisendrath 2016[2]
Country	Italy[1] US[2]
Diagnostic status	DSM-IV-TR MDD (single or recurrent episode), confirmed with MINI[1] DSM-IV unipolar MDD, confirmed with SCID[2]
Age range (mean)	Range NR (mean: 49.0) [1] 16-85 (46.2) [2]
Sex (% female)	72[1] 76[2]
Ethnicity (% BME)	NR[1] 20[2]
Mean age (SD) at first onset of depression	26.9 (12.4) [1] 20.2 (12.2) [2]
Mean months (SD) since onset of current episode	25.5 (47.9) [1] 81.6 (106.8). 59% had chronic depression (>2 years) [2]
No. (SD) of previous depressive episodes	Mean NR (70% ≥3 episodes) [1] 3.7 (2.5) [2]
Details of inadequate response/treatment resistance	Inadequate response (failure to achieve remission, HAMD score≥8) to treatment with antidepressants at adequate dosages for at least 8 weeks before study beginning[1] TRD: Inadequate response to two or more adequate trials prescribed during the current episode assessed with the Antidepressant Treatment History Form (ATHF) [2]
Augmented/previous treatment	Augmented antidepressant: 63% SSRI; 14% SNRIs; 23% other antidepressants[1] Augmented antidepressant: NR (participants in both conditions were encouraged to continue their antidepressant treatment as prescribed by their outside provider) [2]
Baseline severity	HAMD 16.4 (Less severe) [1] HAMD 17.9 (Less severe) [2]
Intervention details (mean dose)	Mindfulness-based cognitive therapy (MBCT; following the manual of Segal et al. 2002) 8x 2-hour weekly sessions[1] Mindfulness-based cognitive therapy (MBCT; adapted from manual by Segal et al. 2002 and based on Chartier et al. 2010) 8x 2.25-hour weekly sessions[2]
Comparator details (mean dose, if applicable)	Attention-placebo (psychoeducational control group) 8x 2-hour weekly sessions[1]

Update 2017

	Mindfulness-based cognitive therapy (MBCT) versus attention-placebo
	Attention-placebo (health enhancement program adapted from manual by MacCoon et al. 2012) 8x 2.25-hour weekly sessions[2]
Treatment length (weeks)	8
Notes: Abbreviations: NR=not reported, SD=standard deviation, TRD=treatment-resistant depression [1]Chiesa 2015; [2]Eisendrath 2016	

Table 118: Summary of findings table for augmenting the antidepressant with a psychological intervention versus attention-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				
	Attention-placebo	Augmenting the antidepressant with a psych intervention				
Remission (Mindfulness-based cognitive therapy [MBCT] versus attention-placebo) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.57** (0.81 to 3.02)	173 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	140 per 1000	**219 per 1000** (113 to 421)				
	Moderate					
	140 per 1000	**220 per 1000** (113 to 423)				
Response (Mindfulness-based cognitive therapy [MBCT] versus attention-placebo) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 2.05** (1.14 to 3.71)	173 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	151 per 1000	**310 per 1000** (172 to 561)				
	Moderate					
	151 per 1000	**310 per 1000** (172 to 560)				
Depression symptomatology (Mindfulness-based cognitive therapy [MBCT] versus attention-placebo) Hamilton Rating Scale for Depression (HAM-D; change score)		The mean depression symptomatology (mindfulness-based cognitive therapy [MBCT] versus attention-placebo) in the intervention groups was **5.06 lower** (7.78 to 2.34 lower)		43 (1 study)	⊕⊕⊖⊖ **low**[5,6]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Attention-placebo	Augmenting the antidepressant with a psych intervention				
Follow-up: mean 8 weeks						
Discontinuation for any reason (Mindfulness-based cognitive therapy [MBCT] versus attention-placebo) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 8 weeks	**Study population**		**RR 0.73** (0.39 to 1.34)	223 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,7]	
	182 per 1000	**133 per 1000** (71 to 244)				
	Moderate					
	206 per 1000	**150 per 1000** (80 to 276)				

Notes:
[1] Unclear method of allocation concealment and non-blind intervention administration
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Data cannot be extracted/is not reported for all outcomes
[4] Events<300
[5] Non-blind intervention administration
[6] N<400
[7] 95% CI crosses line of no effect and both threshold for clinically important benefit (RR 0.75) and for clinically important harm (RR 1.25)

Update 2017

Table 119: Study information table for trials included in the meta-analysis of augmenting the antidepressant with a psychological intervention versus continuing with the antidepressant-only (part 1)

	CBASP + any AD versus any AD	CBT individual (over 15 sessions) + TAU versus TAU	CBT individual (under 15 sessions) + TAU versus TAU
Total no. of studies (N randomised)	1 (491)	2 (627)	1 (42)
Study ID	Kocsis 2009/Klein 2011	Paykel 1999/Scott 2000[1] Wiles 2013/2016[2]	Watkins 2011a
Country	US	UK	UK
Diagnostic status	DSM-IV MDD; chronic depression (depressive symptoms for more than 2 years without remission)	DSM-III-R MDD[1] ICD-10 depressive episode, confirmed with revised clinical interview schedule[2]	DSM-IV major depression (residual symptoms)
Age range (mean)	18-75 (45.4)	21-65 (43.4) [1] Range NR (49.6) [2]	Range NR (44.2)
Sex (% female)	55	49[1] 72[2]	57

	CBASP + any AD versus any AD	CBT individual (over 15 sessions) + TAU versus TAU	CBT individual (under 15 sessions) + TAU versus TAU
Ethnicity (% BME)	11	NR[1] 2[2]	5
Mean age (SD) at first onset of depression	26.4 (13.2)	NR	NR
Mean months (SD) since onset of current episode	92.1 (114.0). 100% chronic depression (MDD≥2 years)	Median: 13.8[1] NR (70% receiving present course of ADs for >12 months)[2]	8.4 (6.2)
No. (SD) of previous depressive episodes	2.6 (3.4)	NR (33% in their first episode)[1] NR (52% ≥5)[2]	5.1 (3.0)
Details of inadequate response/treatment resistance	Inadequate response to 12 weeks of antidepressant medication according to a pharmacotherapy algorithm. Inadequate response defined as failing to meet criteria for remission (≥60% reduction in HAMD score, a HAMD total score<8, and no longer meeting DSM-IV criteria for MDD for 2 consecutive visits during weeks 6 through 12)	Inadequate response (residual symptoms, ≥8 on HAMD and ≥9 on BDI) to antidepressant medication (TCA, SSRI, atypical antidepressant or MAOI) for at least the previous 8 weeks, with at least 4 weeks at an adequate dose, defined as a minimum equivalent to 125mg/day of amitriptyline (and higher levels unless there were definite current side effects or patient refusal to increase dose)[1] Inadequate response (BDI-II≥14) to an adhered to, adequate dose of antidepressant medication (based on BNF and advice from psychopharmacology experts) for at least 6 weeks[2]	Inadequate response (score≥8 on the 17-item Hamilton Depression Rating Scale for Depression [HAMD] and score≥9 on the Beck Depression Inventory [BDI-II]) to antidepressant medication taken at a therapeutic dose as recommended by the British National Formulary and/or equivalent to 125 mg of amitriptyline for at least 8 weeks continuously during the current episode and within the past 2 months
Augmented/previous treatment	Augmented antidepressant: Any AD algorithm-led (began with 2 SSRIs [sertraline and escitalopram], then bupropion [following no response to 2 adequate SSRI trials or to augment treatment in those showing partial SSRI response], then additional options [for	Augmented antidepressant: 60% SSRI (doses equivalent to 33.5mg/day of fluoxetine); 40% TCA (doses equivalent to 186mg/day of amitriptyline)[1] Augmented antidepressant: TAU (participants were taking antidepressants at the time of	Augmented antidepressant: 90% SSRIs/SNRIs; 5% TCAs; 5% MAOIS

Update 2017

	CBASP + any AD versus any AD	CBT individual (over 15 sessions) + TAU versus TAU	CBT individual (under 15 sessions) + TAU versus TAU
	those not benefitting from any of the previous 3] including venlafaxine, mirtazapine, and lithium augmentation)	randomisation and were expected to continue with these drugs as part of their usual care from their general practitioner [SSRIs most common antidepressant taken at baseline: 71%])[2]	
Baseline severity	HAMD 19.3 (Less severe)	HAMD 12.2 (Less severe) [1] BDI 31.8 (More severe) [2]	HAMD 12.7 (Less severe)
Intervention details (mean dose)	Cognitive behavioural analysis system of psychotherapy (CBASP) + any AD (algorithm-based) 16-20 sessions (mean attended 12.5 sessions [SD=6.6])	CBT individual 16 sessions + clinical management (5x 30-min sessions) [1] CBT individual 12x 50-60min sessions with up to a further 6 sessions when judged to be clinically appropriate, maximum of 18 sessions (median number attended 11 sessions) + TAU[2]	Rumination-focused CBT (following methods of Watkins et al. 2007 and Watkins 2009) 12 sessions (mean attended 11 sessions) + TAU (ongoing maintenance antidepressant medication and outpatient clinical management)
Comparator details (mean dose, if applicable)	Any AD (algorithm-led)	Clinical management 5x 30-min sessions[1] TAU (antidepressant treatment and clinical management from GP)[2]	TAU (ongoing maintenance antidepressant medication [90% SSRIs/SNRIs; 5% TCAs; 5% MAOIS], outpatient clinical management and 33% commenced psychological treatment during the trial)
Treatment length (weeks)	12	20[1] 27[2]	12-24 weeks
Notes: Abbreviations: AD=antidepressant, mg=milligrams, NR=not reported, SD=standard deviation [1]Paykel 1999/Scott 2000; [2]Wiles 2013/2016 Note that Kocsis 2009/Klein 2011 is a three-armed trials and demographics reported here are for all three arms combined			

Update 2017

Table 120: Study information table for trials included in the meta-analysis of augmenting the antidepressant with a psychological intervention versus continuing with the antidepressant-only (part 2)

	IPT + TAU versus TAU	Short-term psychodynamic psychotherapy individual + any AD/TAU versus any AD/TAU	Long-term psychodynamic psychotherapy + TAU versus TAU
Total no. of studies (N randomised)	1 (40)	2 (551)	1 (129)
Study ID	Souza 2016	Kocsis 2009/Klein 2011[1] Town (unpublished)[2]	Fonagy 2015
Country	Brazil	US[1] Canada[2]	UK
Diagnostic status	DSM-IV MDD, confirmed with MINI	DSM-IV MDD; chronic depression (depressive symptoms for more than 2 years without remission) [1] DSM-IV MDD, confirmed with MINI and SCID[2]	DSM-IV MDD, confirmed with SCID. Chronic depression (minimum duration of two years of the current depressive episode)
Age range (mean)	Range NR (49.2)	18-75 (45.4) [1] Range NR (41.6) [2]	Range NR (44.3)
Sex (% female)	85	55[1] 63[2]	66
Ethnicity (% BME)	NR	11[1] 3[2]	NR
Mean age (SD) at first onset of depression	35.7 (16.2)	26.4 (13.2) [1] NR[2]	NR
Mean months (SD) since onset of current episode	30.9 (31.3)	92.1 (114.0). 100% chronic depression (MDD≥2 years) [1] Median: 30[2]	45.0 (36.4). 100% had chronic depression (MDD≥2 years)
No. (SD) of previous depressive episodes	2.5 (1.8)	2.6 (3.4) [1] 3.9 (1.6) [2]	NR
Details of inadequate response/treatment resistance	Inadequate response to one trial of antidepressant medication in adequate dose (defined as the equivalent of at least 75mg of amitriptyline) and duration (at least 4 weeks). Participants were under this antidepressant scheme at the moment of randomization	Inadequate response to 12 weeks of antidepressant medication according to a pharmacotherapy algorithm. Inadequate response defined as failing to meet criteria for remission (≥60% reduction in HAMD score, a HAMD total score<8, and no longer meeting DSM-IV criteria for MDD for 2 consecutive visits during weeks 6 through 12) [1]	TRD: Inadequate response to least two different treatments (mean of 3.7 previously failed treatment attempts)

Update 2017

	IPT + TAU versus TAU	Short-term psychodynamic psychotherapy individual + any AD/TAU versus any AD/TAU	Long-term psychodynamic psychotherapy + TAU versus TAU
		Inadequate response (HAMD score≥16) to at least 1 course of antidepressant treatment at the adequate recommended therapeutic dose (34% two or more failed antidepressants for current episode) [2]	
Augmented/previous treatment	Augmented antidepressant: TAU (pharmacotherapy freely chosen by the clinician)	Augmented antidepressant: Any AD algorithm-led (began with 2 SSRIs [sertraline and escitalopram], then bupropion [following no response to 2 adequate SSRI trials or to augment treatment in those showing partial SSRI response], then additional options [for those not benefitting from any of the previous 3] including venlafaxine, mirtazapine, and lithium augmentation) [1] Augmented antidepressant: NR (primary care medication management)[2]	Augmented antidepressant: TAU (82% antidepressants; 41% anxiolytics/hypnotics; 12% antipsychotics/mood stabilizers; 39% analgesics; 29% other medications; 7% no medication; 10% CBT; 14% counselling)
Baseline severity	HAMD 19 (Less severe)	HAMD 19.3 (Less severe) [1] HAMD 23.8 (Less severe) [2]	HAMD 20.1 (Less severe)
Intervention details (mean dose)	Interpersonal Psychotherapy (IPT) 16x 40-min weekly sessions (mean number attended 11.53 sessions) + TAU (pharmacotherapy [freely chosen by the clinician] + clinical management 4-5 sessions [mean attended 4.53])	Brief Supportive Psychotherapy 16-20 sessions (mean attended 13.1 sessions [SD=7.0]) + any AD (algorithm-based) [1] Intensive Short-Term Dynamic Psychotherapy (ISTDP) 20 sessions (mean number attended 16.1	Long-term psychodynamic psychotherapy (following manual by Taylor 2015) 60x 50-min weekly sessions (mean received 41.4 hours [SD=21.4]) + TAU (85% antidepressants; 40% anxiolytics/hypnotics; 12% antipsychotics/mood

Update 2017

	IPT + TAU versus TAU	Short-term psychodynamic psychotherapy individual + any AD/TAU versus any AD/TAU	Long-term psychodynamic psychotherapy + TAU versus TAU
		sessions [SD=6.68]) + TAU (primary care medication management) [2]	stabilizers; 36% analgesics; 24% other medications; 8% no medication; 2% CBT; 2% counselling; 5% self-help groups)
Comparator details (mean dose, if applicable)	TAU (pharmacotherapy [freely chosen by the clinician] + clinical management 4-5 sessions [mean attended 4.27])	Any AD (algorithm-led)[1] TAU (secondary care TAU consisted of a multidisciplinary team approach including pharmacotherapy and clinical management. 97% received at least one session of talking therapy [counselling or CBT]) [2]	TAU (79% antidepressants; 42% anxiolytics/hypnotics; 11% antipsychotics/mood stabilizers; 42% analgesics; 34% other medications; 6% no medication; 19% CBT; 27% counselling; 5% self-help groups)
Treatment length (weeks)	19	12[1] 20[2]	78
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression [1]Kocsis 2009/Klein 2011; [2]Town (unpublished) Note that Kocsis 2009/Klein 2011[1] is a three-armed trials and demographics reported here are for all three arms combined			

Update 2017

Table 121: Study information table for trials included in the meta-analysis of augmenting the antidepressant with a psychological intervention versus continuing with the antidepressant-only (part 3)

	Cognitive and cognitive behavioural therapies (combined) + any AD/TAU versus any AD/TAU-only	Cognitive bibliotherapy + any AD versus any AD	Mutual peer support + TAU versus TAU
Total no. of studies (N randomised)	4 (1160)	1 (90)	1 (463)
Study ID	Kocsis 2009/Klein 2011[1] Paykel 1999/Scott 2000[2] Watkins 2011a[3] Wiles 2013/2016[4]	Schlogelhofer 2014	Valenstein 2016
Country	US[1] UK[2,3,4]	Austria	US
Diagnostic status	DSM-IV MDD; chronic depression (depressive symptoms	DSM-IV-TR MDD	Clinical diagnosis of depression (provider coded a depression diagnosis and

	Cognitive and cognitive behavioural therapies (combined) + any AD/TAU versus any AD/TAU-only	Cognitive bibliotherapy + any AD versus any AD	Mutual peer support + TAU versus TAU
	for more than 2 years without remission) [1] DSM-III-R MDD[2] DSM-IV major depression (residual symptoms) [3] ICD-10 depressive episode, confirmed with revised clinical interview schedule[4]		confirmed that depression was the working diagnosis)
Age range (mean)	18-75 (45.4) [1] 21-65 (43.4) [2] Range NR (44.2) [3] Range NR (49.6) [4]	Range NR (47.8)	Range NR (54.9)
Sex (% female)	55[1] 49[2] 57[3] 72[4]	67	19
Ethnicity (% BME)	11[1] NR[2] 5[3] 2[4]	NR	24
Mean age (SD) at first onset of depression	26.4 (13.2) [1] NR[2,3,4]	NR	NR
Mean months (SD) since onset of current episode	92.1 (114.0). 100% chronic depression (MDD≥2 years) [1] Median: 13.8[2] 8.4 (6.2) [3] NR (70% receiving present course of ADs for >12 months) [4]	NR	NR
No. (SD) of previous depressive episodes	2.6 (3.4) [1] NR (33% in their first episode) [2] 5.1 (3.0) [3] NR (52% ≥5) [4]	NR	NR
Details of inadequate response/treatment resistance	Inadequate response to 12 weeks of antidepressant medication according to a pharmacotherapy algorithm. Inadequate response defined as failing to meet criteria for remission (≥60% reduction in HAMD score, a HAMD total score<8, and no	Inadequate response (not achieving full remission, HAMD score 10-19) to at least one course of a recommended dose of an antidepressant medication for at least 4 weeks (the median treatment duration with antidepressant medication before	Inadequate response (PHQ-9≥10) to at least one prior antidepressant or psychotherapy trial (in the year prior to enrolment 91% received an antidepressant)

Update 2017

	Cognitive and cognitive behavioural therapies (combined) + any AD/TAU versus any AD/TAU-only	Cognitive bibliotherapy + any AD versus any AD	Mutual peer support + TAU versus TAU
	longer meeting DSM-IV criteria for MDD for 2 consecutive visits during weeks 6 through 12) [1] Inadequate response (residual symptoms, ≥8 on HAMD and ≥9 on BDI) to antidepressant medication (TCA, SSRI, atypical antidepressant or MAOI) for at least the previous 8 weeks, with at least 4 weeks at an adequate dose, defined as a minimum equivalent to 125mg/day of amitriptyline (and higher levels unless there were definite current side effects or patient refusal to increase dose) [2] Inadequate response (score≥8 on the 17-item Hamilton Depression Rating Scale for Depression [HAMD] and score≥9 on the Beck Depression Inventory [BDI-II]) to antidepressant medication taken at a therapeutic dose as recommended by the British National Formulary and/or equivalent to 125 mg of amitriptyline for at least 8 weeks continuously during the current episode and within the past 2 months[3] Inadequate response (BDI-II≥14) to an adhered to, adequate dose of antidepressant medication (based on BNF and advice from	screening was 6 months)	

Update 2017

	Cognitive and cognitive behavioural therapies (combined) + any AD/TAU versus any AD/TAU-only	Cognitive bibliotherapy + any AD versus any AD	Mutual peer support + TAU versus TAU
	psychopharmacology experts) for at least 6 weeks[4]		
Augmented/previous treatment	Augmented antidepressant: Any AD algorithm-led (began with 2 SSRIs [sertraline and escitalopram], then bupropion [following no response to 2 adequate SSRI trials or to augment treatment in those showing partial SSRI response], then additional options [for those not benefitting from any of the previous 3] including venlafaxine, mirtazapine, and lithium augmentation)[1] Augmented antidepressant: 60% SSRI (doses equivalent to 33.5mg/day of fluoxetine); 40% TCA (doses equivalent to 186mg/day of amitriptyline)[2] Augmented antidepressant: 90% SSRIs/SNRIs; 5% TCAs; 5% MAOIS[3] Augmented antidepressant: TAU (participants were taking antidepressants at the time of randomisation and were expected to continue with these drugs as part of their usual care from their general practitioner [SSRIs most common antidepressant taken at baseline: 71%])[4]	Augmented antidepressant: NR (all participants were treated with one or more antidepressant drug in clinically adequate doses before and during the trial)	Augmented antidepressant: NR (TAU; 91% antidepressant)
Baseline severity	HAMD 19.3 (Less severe)[1]	HAMD 12.6 (Less severe)	BDI-II 25.4 (More severe)

Update 2017

	Cognitive and cognitive behavioural therapies (combined) + any AD/TAU versus any AD/TAU-only	Cognitive bibliotherapy + any AD versus any AD	Mutual peer support + TAU versus TAU
	HAMD 12.2 (Less severe) [2] HAMD 12.7 (Less severe) [3] BDI 31.8 (More severe) [4]		
Intervention details (mean dose)	Cognitive behavioural analysis system of psychotherapy (CBASP) + any AD (algorithm-based) 16-20 sessions (mean attended 12.5 sessions [SD=6.6]) [1] CBT individual 16 sessions + clinical management (5x 30-min sessions) [2] Rumination-focused CBT (following methods of Watkins et al. 2007 and Watkins 2009) 12 sessions (mean attended 11 sessions) + TAU (ongoing maintenance antidepressant medication and outpatient clinical management) [3] CBT individual 12x 50-60min sessions with up to a further 6 sessions when judged to be clinically appropriate, maximum of 18 sessions (median number attended 11 sessions) + TAU[4]	Cognitive bibliotherapy with 1 monitoring session + any AD	Peer support intervention-Depression Intervention, Actively Learning and Understanding With Peers (DIAL-UP) 1x 2-3 hour training session for peer partner (mean number of calls between pairs 8.6) + TAU (usual mental health care + self-help materials)
Comparator details (mean dose, if applicable)	Any AD (algorithm-led)[1] Clinical management 5x 30-min sessions[2] TAU (ongoing maintenance antidepressant medication [90% SSRIs/SNRIs; 5% TCAs; 5% MAOIS], outpatient clinical management and 33% commenced	Any AD	TAU (usual mental health care + self-help materials)

Update 2017

	Cognitive and cognitive behavioural therapies (combined) + any AD/TAU versus any AD/TAU-only	Cognitive bibliotherapy + any AD versus any AD	Mutual peer support + TAU versus TAU
	psychological treatment during the trial) [3] TAU (antidepressant treatment and clinical management from GP)[4]		
Treatment length (weeks)	12[1] 20[2] 12-24[3] 27[4]	6	24
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Kocsis 2009/Klein 2011; [2]Paykel 1999/Scott 2000; [3]Watkins 2011a; [4]Wiles 2013/2016 Note that Kocsis 2009/Klein 2011[1] is a three-armed trials and demographics reported here are for all three arms combined			

Table 122: Summary of findings table for augmenting the antidepressant with a psychological intervention versus continuing with the antidepressant-only

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
Remission (CBASP + any AD versus any AD) Number of people scoring <8 on Hamilton Rating Scale for Depression (HAM-D) AND responding (≥50% improvement on HAM-D) Follow-up: mean 12 weeks	**Study population**		**RR 0.98** (0.7 to 1.36)	250 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	395 per 1000	**387 per 1000** (276 to 537)				
	Moderate					
	395 per 1000	**387 per 1000** (276 to 537)				
Remission (CBT individual [over 15 sessions] + TAU versus TAU) Number of people scoring ≤7 on	**Study population**		**RR 1.89** (1.34 to 2.66)	577 (2 studies)	⊕⊖⊖⊖ **very low**[4,5]	
	141 per 1000	**266 per 1000** (189 to 375)				
	Moderate					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	**Augmenting the antidepressant with a psych intervention**				
Hamilton Rating Scale for Depression (HAM-D)/<10 on Beck Depression Inventory (BDI) Follow-up: 20-27 weeks	**133 per 1000**	**251 per 1000** (178 to 354)				
Remission (CBT individual [under 15 sessions] + TAU versus TAU) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean – 24 weeks	**Study population**		**RR 3.25** (1.27 to 8.35)	42 (1 study)	⊕⊕⊖⊖ **low**[5,6]	
	190 per 1000	**619 per 1000** (242 to 1000)				
	Moderate					
	191 per 1000	**621 per 1000** (243 to 1000)				
Remission (cognitive and cognitive behavioural therapies [combined] + any AD/TAU versus any AD/TAU-only) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D)/<8 on HAMD AND responding (≥50% improvement on HAM-D)/<10 on Beck Depression Inventory (BDI) Follow-up: 12-27 weeks	**Study population**		**RR 1.68** (1.02 to 2.78)	869 (4 studies)	⊕⊖⊖⊖ **very low**[1,5,7]	
	193 per 1000	**325 per 1000** (197 to 537)				
	Moderate					
	170 per 1000	**286 per 1000** (173 to 473)				
Remission (IPT + TAU versus TAU) Number of people scoring ≤7 on Hamilton Rating Scale for	**Study population**		**RR 1.88** (0.53 to 6.63)	34 (1 study)	⊕⊖⊖⊖ **very low**[2,8]	
	167 per 1000	**313 per 1000** (88 to 1000)				
	Moderate					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
Depression (HAM-D) Follow-up: mean 19 weeks	**167 per 1000**	**314 per 1000** (89 to 1000)				
Remission (short-term psychodynamic psychotherapy individual + any AD/TAU versus any AD/TAU) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D)/<8 on HAMD-D AND responding (≥50% improvement on HAM-D) Follow-up: 12-20 weeks	**Study population**		**RR 2.5** (0.16 to 39.74)	304 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,9,10]	
	292 per 1000	**731 per 1000** (47 to 1000)				
	Moderate					
	214 per 1000	**535 per 1000** (34 to 1000)				
Remission (long-term psychodynamic psychotherapy + TAU versus TAU) Number of people scoring ≤8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 78 weeks	**Study population**		**RR 1.39** (0.41 to 4.69)	129 (1 study)	⊕⊖⊖⊖ **very low**[2,11,12]	
	65 per 1000	**90 per 1000** (26 to 303)				
	Moderate					
	65 per 1000	**90 per 1000** (27 to 305)				
Response (any psych + TAU versus TAU-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Beck Depression Inventory (BDI) Follow-up: 19-27 weeks	**Study population**		**RR 2.22** (1.7 to 2.9)	495 (3 studies)	⊕⊖⊖⊖ **very low**[4,5]	
	218 per 1000	**485 per 1000** (371 to 633)				
	Moderate					
	222 per 1000	**493 per 1000** (377 to 644)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
Response (CBT individual [over 15 sessions] + TAU versus TAU) Number of people showing ≥50% improvement on Beck Depression Inventory (BDI) Follow-up: mean 27 weeks	**Study population**		**RR 2.14** (1.59 to 2.87)	419 (1 study)	⊕⊖⊖⊖ **very low**[5,13]	
	216 per 1000	**462 per 1000** (343 to 620)				
	Moderate					
	216 per 1000	**462 per 1000** (343 to 620)				
Response (CBT individual [under 15 sessions] + TAU versus TAU) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean - weeks	**Study population**		**RR 3.4** (1.54 to 7.51)	42 (1 study)	⊕⊕⊖⊖ **low**[5,6]	
	238 per 1000	**810 per 1000** (367 to 1000)				
	Moderate					
	238 per 1000	**809 per 1000** (367 to 1000)				
Response (IPT + TAU versus TAU) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 19 weeks	**Study population**		**RR 1.69** (0.58 to 4.92)	34 (1 study)	⊕⊖⊖⊖ **very low**[2,8]	
	222 per 1000	**376 per 1000** (129 to 1000)				
	Moderate					
	222 per 1000	**375 per 1000** (129 to 1000)				
Response (cognitive and cognitive behavioural therapies [combined] + TAU versus TAU-only) Number of people showing ≥50% improvement on Hamilton Rating Scale for	**Study population**		**RR 2.32** (1.64 to 3.27)	461 (2 studies)	⊕⊖⊖⊖ **very low**[4,5]	
	218 per 1000	**506 per 1000** (357 to 713)				
	Moderate					
	227 per 1000	**527 per 1000** (372 to 742)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
Depression (HAM-D)/Beck Depression Inventory (BDI) Follow-up: mean 27 weeks						
Depression symptomatology (CBASP + any AD versus any AD) Hamilton Rating Scale for Depression (HAMD; change score) Follow-up: mean 12 weeks		The mean depression symptomatology (CBASP + any AD versus any AD) in the intervention groups was **0.36 standard deviations lower** (0.64 to 0.09 lower)		250 (1 study)	⊕⊖⊖⊖ **very low**[1,3,14]	SMD -0.36 (-0.64 to -0.09)
Depression symptomatology (CBT individual [over 15 sessions] + clinical management/TAU versus clinical management/TAU) Hamilton Rating Scale for Depression (HAM-D; change score)/Beck Depression Inventory (BDI; change score) Follow-up: 20-27 weeks		The mean depression symptomatology (CBT individual [over 15 sessions] + clinical management/TAU versus clinical management/tau) in the intervention groups was **0.41 standard deviations lower** (0.85 lower to 0.04 higher)		577 (2 studies)	⊕⊖⊖⊖ **very low**[4,9,15]	SMD -0.41 (-0.85 to 0.04)
Depression symptomatology (CBT individual [under 15 sessions] + TAU versus TAU) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: -		The mean depression symptomatology (CBT individual [under 15 sessions] + TAU versus TAU) in the intervention groups was **1.29 standard deviations lower** (1.96 to 0.62 lower)		42 (1 study)	⊕⊕⊖⊖ **low**[6,14]	SMD -1.29 (-1.96 to -0.62)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
Depression symptomatology (IPT + TAU versus TAU) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 19 weeks		The mean depression symptomatology (IPT + TAU versus TAU) in the intervention groups was **0.66 standard deviations lower** (1.35 lower to 0.04 higher)		34 (1 study)	⊕⊕⊖⊖ **low**[8,15]	SMD -0.66 (-1.35 to 0.04)
Depression symptomatology (short-term psychodynamic psychotherapy individual + any AD versus any AD) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 12 weeks		The mean depression symptomatology (short-term psychodynamic psychotherapy individual + any AD versus any AD) in the intervention groups was **0.1 standard deviations lower** (0.37 lower to 0.17 higher)		244 (1 study)	⊕⊖⊖⊖ **very low**[1,3,14]	SMD -0.1 (-0.37 to 0.17)
Depression symptomatology (long-term psychodynamic psychotherapy + TAU versus TAU-only) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 78 weeks		The mean depression symptomatology (long-term psychodynamic psychotherapy + tAU versus TAU-only) in the intervention groups was **0.26 standard deviations lower** (0.61 lower to 0.09 higher)		129 (1 study)	⊕⊖⊖⊖ **very low**[11,12,15]	SMD -0.26 (-0.61 to 0.09)
Depression symptomatology (cognitive bibliotherapy + any AD versus any AD) Hamilton Rating Scale for Depression (HAM-		The mean depression symptomatology (cognitive bibliotherapy + any AD versus any AD) in the intervention		90 (1 study)	⊕⊕⊖⊖ **low**[6,15]	SMD -0.37 (-0.79 to 0.05)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
D; change score) Follow-up: mean 6 weeks		groups was **0.37 standard deviations lower** (0.79 lower to 0.05 higher)				
Depression symptomatology (mutual peer support + TAU versus TAU) Beck Depression Inventory (BDI-II; change score) Follow-up: mean 24 weeks		The mean depression symptomatology (mutual peer support + TAU versus TAU) in the intervention groups was **0.03 standard deviations lower** (0.25 lower to 0.19 higher)		344 (1 study)	⊕⊖⊖⊖ **very low**[1,14,16]	SMD -0.03 (-0.25 to 0.19)
Depression symptomatology (cognitive and cognitive behavioural therapies [combined] + any AD/TAU versus any AD/TAU-only) Hamilton Rating Scale for Depression (HAM-D; change score)/Beck Depression Inventory (BDI; change score) Follow-up: 12-27 weeks		The mean depression symptomatology (cognitive and cognitive behavioural therapies [combined] + any AD/TAU versus any AD/TAU-only) in the intervention groups was **0.52 standard deviations lower** (0.83 to 0.2 lower)		869 (4 studies)	⊕⊕⊖⊖ **low**[7,17]	SMD -0.52 (-0.83 to -0.2)
Discontinuation for any reason (CBASP + any AD versus any AD) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 12 weeks	**Study population**		**RR 0.75** (0.42 to 1.34)	296 (1 study)	⊕⊖⊖⊖ **very low**[1,3,18]	
	167 per 1000	**125 per 1000** (70 to 223)				
	Moderate					
	167 per 1000	**125 per 1000** (70 to 224)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
Discontinuation for any reason (CBT individual [over 15 sessions] + clinical management/TAU versus clinical management/TAU) Number of participants discontinuing for any reason (including adverse events) Follow-up: 20-27 weeks	**Study population**		**RR 1.29** (0.85 to 1.96)	627 (2 studies)	⊕⊕⊖⊖ **low**[1,19]	
	109 per 1000	**140 per 1000** (92 to 213)				
	Moderate					
	124 per 1000	**160 per 1000** (105 to 243)				
Discontinuation for any reason (CBT individual [under 15 sessions] + TAU versus TAU) Number of participants discontinuing for any reason (including adverse events)	**Study population**		**RR 0.5** (0.05 to 5.1)	42 (1 study)	⊕⊖⊖⊖ **very low**[6,18]	
	95 per 1000	**48 per 1000** (5 to 486)				
	Moderate					
	95 per 1000	**48 per 1000** (5 to 484)				
Discontinuation for any reason (IPT + TAU versus TAU) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 19 weeks	**Study population**		**RR 3.38** (0.74 to 15.39)	40 (1 study)	⊕⊖⊖⊖ **very low**[6,18]	
	87 per 1000	**294 per 1000** (64 to 1000)				
	Moderate					
	87 per 1000	**294 per 1000** (64 to 1000)				
Discontinuation for any reason (short-term psychodynamic psychotherapy individual + any	**Study population**		**RR 1.19** (0.45 to 3.13)	351 (2 studies)	⊕⊖⊖⊖ **very low**[1,7,10,18]	
	151 per 1000	**179 per 1000** (68 to 472)				
	Moderate					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
AD/TAU versus any AD/TAU) Number of participants discontinuing for any reason (including adverse events) Follow-up: 12-20 weeks	**133 per 1000**	**158 per 1000** (60 to 416)				
Discontinuation for any reason (long-term psychodynamic psychotherapy + TAU versus TAU-only) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 78 weeks	**Study population**		**RR 1.16** (0.49 to 2.74)	129 (1 study)	⊕⊖⊖⊖ **very low**[11,12,18]	
	129 per 1000	**150 per 1000** (63 to 354)				
	Moderate					
	129 per 1000	**150 per 1000** (63 to 353)				
Discontinuation for any reason (cognitive bibliotherapy + any AD versus any AD) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 6 weeks	**Study population**		**RR 1.53** (0.62 to 3.79)	90 (1 study)	⊕⊖⊖⊖ **very low**[6,18]	
	146 per 1000	**224 per 1000** (91 to 555)				
	Moderate					
	146 per 1000	**223 per 1000** (91 to 553)				
Discontinuation for any reason (mutual peer support + TAU versus TAU) Number of participants discontinuing for any reason (including adverse	**Study population**		**RR 0.97** (0.53 to 1.78)	387 (1 study)	⊕⊖⊖⊖ **very low**[1,16,18]	
	107 per 1000	**104 per 1000** (57 to 190)				
	Moderate					
	107 per 1000	**104 per 1000** (57 to 190)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				
events) Follow-up: mean 24 weeks						
Discontinuation for any reason (cognitive and cognitive behavioural therapies [combined] + any AD/TAU versus any AD/TAU-only) Number of participants discontinuing for any reason (including adverse events) Follow-up: 12-27 weeks	**Study population** **121 per 1000** **Moderate** **125 per 1000**	**128 per 1000** (91 to 180) **132 per 1000** (94 to 186)	**RR 1.06** (0.75 to 1.49)	965 (4 studies)	⊕⊖⊖⊖ **very low**[17,18]	
Discontinuation due to adverse events (CBASP + any AD versus any AD) Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Study population** **21 per 1000** **Moderate** **21 per 1000**	**10 per 1000** (1 to 70) **10 per 1000** (1 to 71)	**RR 0.48** (0.07 to 3.36)	296 (1 study)	⊕⊖⊖⊖ **very low**[1,3,18]	
Discontinuation due to adverse events (short-term psychodynamic psychotherapy individual + any AD versus any AD) Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Study population** **21 per 1000** **Moderate** **21 per 1000**	**5 per 1000** (0 to 56) **5 per 1000** (0 to 56)	**RR 0.25** (0.02 to 2.68)	291 (1 study)	⊕⊖⊖⊖ **very low**[1,3,18]	

Notes:
[1] Method of randomisation was unclear, and non-blind participants and intervention administrator(s)
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and for

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant-only	Augmenting the antidepressant with a psych intervention				

clinically important benefit (RR 1.25)
[3] Drugs were supplied at no cost by pharmaceutical company and authors have financial interests with pharmaceutical companies
[4] High risk of bias associated with randomisation method due to significant difference between groups at baseline, non-blind participants and intervention administrator(s), and unclear blinding of outcome assessment, in studies contributing >50% of weighting in analysis
[5] Events<300
[6] Non-blind participants and intervention administrator(s)
[7] I-squared>50%
[8] Non-blind participants and intervention administrator(s) and potential risk of attrition bias (difference in drop-out between groups>20% but ITT analysis used)
[9] I-squared>80%
[10] Data cannot be extracted or is not reported for all outcomes and/or drugs were supplied at no cost by pharmaceutical company and authors have financial interests with pharmaceutical companies
[11] High risk of bias associated with randomisation method due to significant difference between groups at baseline, non-blind participants and intervention administrator(s)
[12] Study partially funded by the International Psychoanalytic Association
[13] High risk of bias associated with randomisation method due to significant difference between groups at baseline, non-blind participants and intervention administrator(s), and unclear blinding of outcome assessment
[14] N<400
[15] 95% CI crosses both line of no effect and threshold for clinically important benefit (SMD -0.5)
[16] Data cannot be extracted/is not reported for all outcomes
[17] High or unclear risk of randomisation method and participants and intervention administrator(s) were non-blind
[18] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[19] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)

Table 123: Study information table for trials included in the meta-analysis of augmenting the antidepressant with a psychological intervention versus augmenting with a non-antidepressant agent

	CBT individual (under 15 sessions) + AD versus lithium + AD
Total no. of studies (N randomised)	1 (44)
Study ID	Kennedy 2003
Country	Canada
Diagnostic status	DSM-IV MDE, confirmed with SCID
Age range (mean)	Range NR (39.3)
Sex (% female)	55
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	25.4 (13.4)
Mean months (SD) since onset of current episode	28.4 (37.8)
No. (SD) of previous depressive episodes	2.2 (1.4)
Details of inadequate response/treatment resistance	Partial response (score of 8-15 on HAMD-D) to 1 of 4 standard antidepressant medications

Update 2017

	CBT individual (under 15 sessions) + AD versus lithium + AD
	(moclobemide, paroxetine, sertraline, or venlafaxine) to maximum tolerated doses for 8-14 weeks
Augmented/previous treatment	Augmented antidepressant: Moclobemide (300-600mg/day), paroxetine (20-40mg/day), sertraline (50-200mg/day), or venlafaxine (75-225mg/day)
Baseline severity	HAMD 11.9 (Less severe)
Intervention details (mean dose)	CBT individual (12 sessions) + AD
Comparator details (mean dose, if applicable)	Lithium 600-900mg/day + AD
Treatment length (weeks)	8
Notes: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation	

Table 124: Summary of findings table for augmenting the antidepressant with a psychological intervention versus augmenting with a non-antidepressant agent

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmenting with a non-AD agent	Augmenting the antidepressant with a psych intervention				
Remission (CBT individual [under 15 sessions] + AD versus lithium + AD) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 0.68** (0.28 to 1.65)	44 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	381 per 1000	**259 per 1000** (107 to 629)				
	Moderate					
	381 per 1000	**259 per 1000** (107 to 629)				
Depression symptomatology (CBT individual [under 15 sessions] + AD versus lithium + AD) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 8 weeks		The mean depression symptomatology (CBT individual [under 15 sessions] + AD versus lithium + AD) in the intervention groups was **0.7 standard deviations higher** (0.09 to 1.31 higher)		44 (1 study)	⊕⊕⊖⊖ **low**[1,3]	SMD 0.7 (0.09 to 1.31)
	Study population					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmenting with a non-AD agent	Augmenting the antidepressant with a psych intervention				
Discontinuation for any reason (CBT individual [under 15 sessions] + AD versus lithium + AD) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 8 weeks	**286 per 1000**	**260 per 1000** (100 to 686) **Moderate** **286 per 1000**	**260 per 1000** (100 to 686)	**RR 0.91** (0.35 to 2.4)	44 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	
Discontinuation due to adverse events (CBT individual [under 15 sessions] + AD versus lithium + AD) Number of participants discontinuing due to adverse events Follow-up: mean 8 weeks	**Study population** **48 per 1000**	**15 per 1000** (0 to 339) **Moderate** **48 per 1000**	**15 per 1000** (0 to 342)	**RR 0.31** (0.01 to 7.12)	44 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	

Notes:
[1] Unclear method of randomisation and allocation concealment, non-blind participants and intervention administrator(s), and unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] N<400
[4] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 125: Study information table for trials included in the meta-analysis of augmenting the antidepressant with CBASP versus augmenting with 'other' psychological intervention (head-to-head comparisons)

	CBASP + any AD versus short-term psychodynamic psychotherapy individual + any AD
Total no. of studies (N randomised)	1 (491)
Study ID	Kocsis 2009/Klein 2011
Country	US

Update 2017

	CBASP + any AD versus short-term psychodynamic psychotherapy individual + any AD
Diagnostic status	DSM-IV MDD; chronic depression (depressive symptoms for more than 2 years without remission)
Age range (mean)	18-75 (45.4)
Sex (% female)	55
Ethnicity (% BME)	11
Mean age (SD) at first onset of depression	26.4 (13.2)
Mean months (SD) since onset of current episode	92.1 (114.0). 100% chronic depression (MDD≥2 years)
No. (SD) of previous depressive episodes	2.6 (3.4)
Details of inadequate response/treatment resistance	Inadequate response to 12 weeks of antidepressant medication according to a pharmacotherapy algorithm. Inadequate response defined as failing to meet criteria for remission (≥60% reduction in HAMD score, a HAMD total score<8, and no longer meeting DSM-IV criteria for MDD for 2 consecutive visits during weeks 6 through 12)
Augmented/previous treatment	Augmented antidepressant: Any AD algorithm-led (began with 2 SSRIs [sertraline and escitalopram], then bupropion [following no response to 2 adequate SSRI trials or to augment treatment in those showing partial SSRI response], then additional options [for those not benefitting from any of the previous 3] including venlafaxine, mirtazapine, and lithium augmentation)
Baseline severity	HAMD 19.3 (Less severe)
Intervention details (mean dose)	Cognitive behavioural analysis system of psychotherapy (CBASP) 16-20 sessions (mean attended 12.5 sessions [SD=6.6]) + any AD (algorithm-based)
Comparator details (mean dose, if applicable)	Brief Supportive Psychotherapy 16-20 sessions (mean attended 13.1 sessions [SD=7.0]) + any AD (algorithm-based)
Treatment length (weeks)	12
Notes: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation Note that Kocsis 2009/Klein 2011 is a three-armed trials and demographics reported here are for all three arms combined	

Update 2017

Table 126: Summary of findings table for augmenting the antidepressant with CBASP versus augmenting with 'other' psychological intervention (head-to-head comparisons)

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	Augmenting the antidepressant with a psych intervention [head-to-head]				
Remission (CBASP + any AD versus short-term psychodynamic psychotherapy individual + any AD) Number of people scoring <8 on Hamilton Rating Scale for Depression (HAM-D) AND responding (≥50% improvement on HAM-D) Follow-up: mean 12 weeks	**Study population**		**RR 1.24** (0.93 to 1.67)	342 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
	310 per 1000	**384 per 1000** (288 to 517)				
	Moderate					
	310 per 1000	**384 per 1000** (288 to 518)				
Depression symptomatology (CBASP + any AD versus short-term psychodynamic psychotherapy individual + any AD) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 12 weeks		The mean depression symptomatology (CBASP + any AD versus short-term psychodynamic psychotherapy individual + any AD) in the intervention groups was **0.26 standard deviations lower** (0.48 to 0.05 lower)		342 (1 study)	⊕⊕⊖⊖ **low**[1,3]	SMD -0.26 (-0.48 to -0.05)
Discontinuation for any reason (CBASP + any AD versus short-term psychodynamic psychotherapy individual + any AD) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 12 weeks	**Study population**		**RR 0.9** (0.54 to 1.5)	395 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	
	138 per 1000	**125 per 1000** (75 to 208)				
	Moderate					
	139 per 1000	**125 per 1000** (75 to 208)				
Discontinuation due to adverse events (CBASP + any AD versus short-term	**Study population**		**RR 1.95** (0.18 to 21.33)	395 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	
	5 per 1000	**10 per 1000** (1 to 109)				

Update 2017

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	Augmenting the antidepressant with a psych intervention [head-to-head]				
psychodynamic psychotherapy individual + any AD) Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Moderate**					
	5 per 1000	**10 per 1000** (1 to 107)				

Notes:
[1] Method of randomisation was unclear, and non-blind participants and intervention administrator(s)
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] N<400
[4] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 127: Study information table for trials included in the meta-analysis of augmenting the antidepressant with a physical intervention versus attention-placebo

	Exercise + SSRI/any AD versus attention-placebo + SSRI/any AD
Total no. of studies (N randomised)	2 (106)
Study ID	Lavretsky 2011[1] Mota-Pereira 2011[2]
Country	US[1] Portugal[2]
Diagnostic status	DSM-IV MDD, confirmed with SCID[1] DSM-IV MDD[2]
Age range (mean)	>60 (70.6) [1] 26-60 (47.5) [2]
Sex (% female)	62[1] 66[2]
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	44.1 (24.1) [1] NR[2]
Mean months (SD) since onset of current episode	35.3 (33.6) [1] NR[2]
No. (SD) of previous depressive episodes	3.8 (4.1) [1] NR[2]
Details of inadequate response/treatment resistance	Inadequate response to 4 weeks prospective treatment with escitalopram[1] Inadequate response (failure to show clinical remission, HAMD>7) to combined therapy in doses considered adequate for 9-15 months[2]

Update 2017

	Exercise + SSRI/any AD versus attention-placebo + SSRI/any AD
Augmented/previous treatment	Augmented antidepressant: Escitalopram (10-20mg/day) [1] Augmented antidepressant: Usual pharmacological therapy (all patients were medicated with non-sedating antidepressants in doses considered therapeutic: clomipramine, maprotiline and amitriptyline were used as tricyclic antidepressants at a dose of 125-150 mg/day; as SSRIs fluoxetine, escitalopram, paroxetine and sertraline were used, at doses of 20-40 mg/day, 20 mg/day, 20-40 mg/day and 100-150 mg/day, respectively; venlafaxine was used as SNRI at a dose of 150 mg/day; when considered appropriate, lorazepam was used as anxiolytic at a dose of 1-2.5 mg/day) [2]
Baseline severity	HAMD 9 (Less severe) [1] HAMD 17 (Less severe) [2]
Intervention details (mean dose)	Tai Chi Chih 10x 2-hour sessions + escitalopram 10-20mg/day (mean dose 12.5 mg/day) [1] Aerobic exercise 60 sessions/12x 30-45min sessions supervised + any AD (usual pharmacological therapy) [2]
Comparator details (mean dose, if applicable)	Attention-placebo (health education) 10x 2-hour sessions + escitalopram 10-20mg/day (mean dose 12.7 mg/day) [1] Attention-placebo 12x 30-45min sessions + any AD (usual pharmacological therapy) [2]
Treatment length (weeks)	10[1] 12[2]
Notes: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation	

Update 2017

Table 128: Summary of findings table for augmenting the antidepressant with a physical intervention versus attention-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				
	Attention-placebo	Augmenting the antidepressant with a physical intervention				
Remission (exercise + SSRI/any AD versus attention-placebo + SSRI/any AD) Number of people scoring ≤7/10 on Hamilton Rating Scale for Depression (HAM-D)	**Study population**		**RR 1.77** (0.37 to 8.41)	102 (2 studies)	⊕⊖⊖⊖ **very low**[1,2]	
	596 per 1000	**1000 per 1000** (220 to 1000)				
	Moderate					
	378 per 1000	**669 per 1000** (140 to 1000)				

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Attention-placebo	Augmenting the antidepressant with a physical intervention				
Follow-up: 10-12 weeks						
Response (exercise + any AD versus attention-placebo + any AD) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 12 weeks	**Study population** **0 per 1000**	**0 per 1000** (0 to 0)	**RR 4.95** (0.29 to 83.68)	29 (1 study)	⊕⊖⊖⊖ **very low**[2,3,4]	
	Moderate **0 per 1000**	**0 per 1000** (0 to 0)				
Depression symptomatology (exercise + SSRI/any AD versus attention-placebo + SSRI/any AD) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: 10-12 weeks		The mean depression symptomatology (exercise + SSRI/any AD versus attention-placebo + SSRI/any AD) in the intervention groups was **0.4 standard deviations lower** (0.86 lower to 0.06 higher)		97 (2 studies)	⊕⊖⊖⊖ **very low**[1,5,6]	SMD -0.4 (-0.86 to 0.06)
Discontinuation for any reason (exercise + SSRI/any AD versus attention-placebo + SSRI/any AD) Number of participants discontinuing for any reason (including adverse events) Follow-up: 10-12 weeks	**Study population** **62 per 1000**	**96 per 1000** (25 to 366)	**RR 1.53** (0.4 to 5.86)	106 (2 studies)	⊕⊖⊖⊖ **very low**[1,7]	
	Moderate **73 per 1000**	**112 per 1000** (29 to 428)				

Notes:
[1] Non-blind intervention administration
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] High risk of bias associated with randomisation method due to significant difference between groups at baseline and unclear method of allocation concealment. Intervention administration was non-blind
[4] Study partially funded by pharmaceutical company
[5] I-squared>80%

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Attention-placebo	Augmenting the antidepressant with a physical intervention				

[6] 95% CI crosses both line of no effect and threshold for clinically important benefit (SMD -0.5)
[7] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

8.3.31 Switching strategies

Evidence was found for seven switching treatment strategy comparisons as follows: switching to another antidepressant of a different class compared to placebo (see Table 129 for study characteristics); switching to another antidepressant of a different class compared to continuing with the antidepressant (see Table 131 for study characteristics); switching to a non-antidepressant agent compared to continuing with the antidepressant (see Table 133 for study characteristics); switching to another antidepressant/non-antidepressant agent compared to augmentation with another antidepressant/non-antidepressant agent (see Table 135 for study characteristics); switching to another antidepressant of the same class compared to switching to another antidepressant of a different class (see Table 137 for study characteristics); head-to-head comparisons of switching to another antidepressant/non-antidepressant agent (see Table 139 and Table 140 for study characteristics); switching to a combined psychological and pharmacological intervention compared to switching to psychological intervention-only (see Table 142 for study characteristics).

Evidence for these comparisons are summarised in the clinical GRADE evidence profiles below (see Table 130, Table 132, Table 134, Table 136, Table 138, Table 141 and Table 143). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J5.

Table 129: Study information table for trials included in the meta-analysis of switching to another antidepressant of a different class versus placebo

	Switch from SSRI to atypical antidepressant or placebo
Total no. of studies (N randomised)	1 (325)
Study ID	GlaxoSmithKline 2009
Country	Japan
Diagnostic status	DSM-IV-TR MDD (single episode or recurrent), without psychotic features
Age range (mean)	Range NR (36.4)
Sex (% female)	45
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Details of inadequate response/treatment resistance	Inadequate response to paroxetine (20-40 mg/day) for 4 weeks
Augmented/previous treatment	Previous treatment: Paroxetine (20-40mg/day)

Update 2017

	Switch from SSRI to atypical antidepressant or placebo
Baseline severity	HAMD 19.6 (Less severe)
Intervention details (mean dose)	Bupropion Hydrochloride Sustained Release (323U66) 100-300mg/day
Comparator details (mean dose, if applicable)	Placebo
Treatment length (weeks)	12
Note: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation	

Table 130: Summary of findings table for switching to another antidepressant of a different class 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				
	Placebo	Switch to another antidepressant of different class				
Remission (SSRI to atypical antidepressant or placebo) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 12 weeks	**Study population**		**RR 0.98** (0.67 to 1.43)	322 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	248 per 1000	**243 per 1000** (166 to 355)				
	Moderate					
	248 per 1000	**243 per 1000** (166 to 355)				
Response (SSRI to atypical antidepressant or placebo) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 12 weeks	**Study population**		**RR 1.03** (0.78 to 1.37)	322 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	369 per 1000	**381 per 1000** (288 to 506)				
	Moderate					
	369 per 1000	**380 per 1000** (288 to 506)				
Response (SSRI to atypical antidepressant or placebo) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 12 weeks	**Study population**		**RR 1.09** (0.86 to 1.38)	322 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	439 per 1000	**479 per 1000** (378 to 606)				
	Moderate					
	440 per 1000	**480 per 1000** (378 to 607)				

Update 2017

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	Switch to another antidepressant of different class				
Depression symptomatology (SSRI to atypical antidepressant or placebo) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 12 weeks		The mean depression symptomatology (SSRI to atypical antidepressant or placebo) in the intervention groups was **0.02 standard deviations higher** (0.19 lower to 0.24 higher)		322 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	SMD 0.02 (-0.19 to 0.24)
Discontinuation for any reason (SSRI to atypical antidepressant or placebo) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 12 weeks	**Study population**		**RR 1.37** (1.01 to 1.85)	325 (1 study)	⊕⊖⊖⊖ **very low**[1,3,6]	
	296 per 1000	**405 per 1000** (299 to 547)				
	Moderate					
	296 per 1000	**406 per 1000** (299 to 548)				
Discontinuation due to adverse events (SSRI to atypical antidepressant or placebo) Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Study population**		**RR 1.21** (0.79 to 1.83)	325 (1 study)	⊕⊖⊖⊖ **very low**[1,3,7]	
	195 per 1000	**236 per 1000** (154 to 357)				
	Moderate					
	195 per 1000	**236 per 1000** (154 to 357)				

Notes:
[1] Unclear randomisation method and method of allocation concealment, and unclear risk of attrition bias (drop-out>20% but difference between groups <20% and ITT analysis used)
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] Study run and funded by pharmaceutical company
[4] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[5] N<400
[6] Events<300
[7] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)

Update 2017

Table 131: Study information table for trials included in the meta-analysis of switching to another antidepressant of a different class versus continuing with the same antidepressant

	Switch to SSRI versus continuing TCA/SNRI	Switch to atypical AD/SNRI/TeCA (mianserin) versus continuing SSRI
Total no. of studies (N randomised)	2 (983)	2 (479)
Study ID	Corya 2006[1] Shelton 2005[2]	Fang 2010/2011[3] Ferreri 2001[4]
Country	16 countries[1] US and Canada[2]	China[3] France[4]
Diagnostic status	DSM-IV MDD (single episode or recurrent), without psychotic features[1] DSM-IV unipolar nonpsychotic MDD, confirmed with the SCID[2]	DSM-IV MDD[3] DSM-III-R MDD[4]
Age range (mean)	Range NR (45.7) [1] Range NR (42.4) [2]	NR[3] Range NR (46.6) [4]
Sex (% female)	73[1] 68[2]	NR[3] 74[4]
Ethnicity (% BME)	10[1] 12[2]	NR
Mean age (SD) at first onset of depression	NR	NR
Mean months (SD) since onset of current episode	Median 26.6[1] Median: 11.8[2]	NR[3] 7.3 (8.4) [4]
No. (SD) of previous depressive episodes	Mean NR (51% >3 episodes) [1] NR[2]	NR[3] 2.4 (2.2) [4]
Details of inadequate response/treatment resistance	TRD: Inadequate response to a selective serotonin reuptake inhibitor (SSRI) antidepressant after at least 6 weeks of therapy at a therapeutic dose (i.e., citalopram, 40 mg/day; fluoxetine, 40 mg/day; paroxetine, 40 mg/day; or sertraline, 150 mg/day) at entry into the trial and inadequate response (<30% improvement in MADRS total score) to an open-label, 7-week lead-in phase of venlafaxine (75–375 mg/day according to the investigator's clinical judgment)[1] TRD: History of at least one failure to respond to SSRI after at least 4 weeks of therapy at a therapeutic dose (i.e. citalopram 40mg/day, fluoxetine 40mg/day, paroxetine 40mg/day, or sertraline 150mg/day), and failure to respond (<30%	TRD: Inadequate response to 2 or more adequate treatments from different classes of antidepressants in the current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment[3] Inadequate response to previous fluoxetine treatment after at least 6 weeks of treatment with 20 mg/day[4]

Update 2017

	Switch to SSRI versus continuing TCA/SNRI	Switch to atypical AD/SNRI/TeCA (mianserin) versus continuing SSRI
	improvement on MADRS) to nortriptyline (25-175mg/day; mean modal dose 104.6mg/day) during a 7-week open-label treatment phase[2]	
Augmented/previous treatment	Previous treatment: Venlafaxine (75–375 mg/day). Mean number of prior psychotropic medications: 4.1[1] Previous treatment: Nortriptyline (25-175mg/day; mean modal dose 104.6mg/day)[2]	Previous treatment: Paroxetine (20mg/day)[3] Previous treatment: Fluoxetine (20mg/day)[4]
Baseline severity	MADRS 30 (More severe)[1] MADRS 28.5 (More severe)[2]	NR[3] HAMD 27.2 (More severe)[4]
Intervention details (mean dose)	Fluoxetine 25 or 50mg/day (mean modal dose 37.5 mg/day)[1] Fluoxetine 25-50mg/day (mean modal dose 35.8mg/day)[2]	Two groups combined: Mirtazapine 45mg/day or Venlafaxine-XR 225mg/day[3] Mianserin 60 mg/day[4]
Comparator details (mean dose, if applicable)	Venlafaxine 75-375mg/day (mean modal dose 275.4 mg/day)[1] Nortriptyline 25-175mg/day (mean modal dose 103.5mg/day)[2]	Paroxetine 20mg/day[3] Fluoxetine 20mg/day[4]
Treatment length (weeks)	12[1] 8[2]	8[3] 6[4]
Notes: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression [1]Corya 2006; [2]Shelton 2005 Note that Corya 2006[1] is a five-armed trial, Fang 2010/2011[3] is an eight-armed trial, Ferreri 2001[4] is a three-armed trial and Shelton 2005[2] is a four-armed trial and demographics reported here are for all arms combined		

Update 2017

Table 132: Summary of findings table for switching to another antidepressant of a different class versus continuing with the same antidepressant

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to another antidepressant of a different class				
Remission (any switch versus continuing with the antidepressant)	**Study population**		**RR 0.93** (0.65 to 1.34)	545 (4 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	254 per 1000	**236 per 1000** (165 to 340)				

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to another antidepressant of a different class				
Number of people scoring ≤7/8 on Hamilton Rating Scale for Depression (HAM-D)/≤8 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 6-12 weeks	**Moderate**					
	204 per 1000	**190 per 1000** (133 to 273)				
Remission (switch to SSRI versus continuing TCA/SNRI) Number of people scoring ≤8 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 0.78** (0.47 to 1.27)	324 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,4]	
	198 per 1000	**155 per 1000** (93 to 252)				
	Moderate					
	200 per 1000	**156 per 1000** (94 to 254)				
Remission (switch to atypical AD/SNRI/TeCA [mianserin] versus continuing SSRI) Number of people scoring ≤7/8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 6-8 weeks	**Study population**		**RR 1.19** (0.52 to 2.77)	221 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,5,6]	
	337 per 1000	**401 per 1000** (175 to 934)				
	Moderate					
	325 per 1000	**387 per 1000** (169 to 900)				
Response (any switch versus continuing with the antidepressant) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Montgomery Asberg Depression Rating Scale (MADRS)	**Study population**		**RR 0.91** (0.74 to 1.12)	545 (4 studies)	⊕⊖⊖⊖ **very low**[1,3,7]	
	450 per 1000	**409 per 1000** (333 to 504)				
	Moderate					
	434 per 1000	**395 per 1000** (321 to 486)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to another antidepressant of a different class				
Follow-up: 6-12 weeks						
Response (switch to SSRI versus continuing TCA/SNRI) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 0.8** (0.58 to 1.09)	324 (2 studies)	⊕⊖⊖⊖ **very low**[3,4,7]	
	397 per 1000	**317 per 1000** (230 to 433)				
	Moderate					
	404 per 1000	**323 per 1000** (234 to 440)				
Response (switch to atypical AD/SNRI/TeCA [mianserin] versus continuing SSRI) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 6-8 weeks	**Study population**		**RR 1.01** (0.73 to 1.41)	221 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,5]	
	530 per 1000	**535 per 1000** (387 to 747)				
	Moderate					
	518 per 1000	**523 per 1000** (378 to 730)				
Response (switch to TeCA [mianserin] versus continuing SSRI) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 6 weeks	**Study population**		**RR 1.42** (0.92 to 2.2)	71 (1 study)	⊕⊖⊖⊖ **very low**[3,8,9]	
	447 per 1000	**635 per 1000** (412 to 984)				
	Moderate					
	447 per 1000	**635 per 1000** (411 to 983)				
Depression symptomatology (any switch versus continuing with the antidepressant) Hamilton Rating Scale for Depression		The mean depression symptomatology (any switch versus continuing with the antidepressant)		400 (3 studies)	⊕⊕⊖⊖ **low**[3,10]	SMD -0.04 (-0.3 to 0.23)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	**Switch to another antidepressant of a different class**				
(HAM-D; change score)/Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 6-12 weeks		in the intervention groups was **0.04 standard deviations lower** (0.3 lower to 0.23 higher)				
Depression symptomatology (switch to SSRI versus continuing TCA/SNRI) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 8-12 weeks		The mean depression symptomatology (switch to SSRI versus continuing TCA/SNRI) in the intervention groups was **0.03 standard deviations higher** (0.31 lower to 0.38 higher)		329 (2 studies)	⊕⊖⊖⊖ **very low**[3,4,6,11]	SMD 0.03 (-0.31 to 0.38)
Depression symptomatology (switch to TeCA [mianserin] versus continuing SSRI) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 6 weeks		The mean depression symptomatology (switch to TeCA [mianserin] versus continuing SSRI) in the intervention groups was **0.24 standard deviations lower** (0.71 lower to 0.23 higher)		71 (1 study)	⊕⊖⊖⊖ **very low**[3,8,12]	SMD -0.24 (-0.71 to 0.23)
Discontinuation for any reason (any switch versus continuing with the antidepressant) Number of participants discontinuing for any reason (including adverse events) Follow-up: 6-12 weeks	**Study population** **181 per 1000** **Moderate** **181 per 1000**	**223 per 1000** (147 to 337) **223 per 1000** (147 to 337)	**RR 1.23** (0.81 to 1.86)	551 (4 studies)	⊕⊖⊖⊖ **very low**[3,13,14]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to another antidepressant of a different class				
Discontinuation for any reason (switch to SSRI versus continuing TCA/SNRI) Number of participants discontinuing for any reason (including adverse events) Follow-up: 8-12 weeks	**Study population**		**RR 1.13** (0.54 to 2.38)	329 (2 studies)	⊕⊖⊖⊖ **very low**[3,6,15,16]	
	181 per 1000	**205 per 1000** (98 to 431)				
	Moderate					
	186 per 1000	**210 per 1000** (100 to 443)				
Discontinuation for any reason (switch to atypical AD/SNRI/TeCA [mianserin] versus continuing SSRI) Number of participants discontinuing for any reason (including adverse events) Follow-up: 6-8 weeks	**Study population**		**RR 1.37** (0.74 to 2.54)	222 (2 studies)	⊕⊖⊖⊖ **very low**[3,16,17]	
	181 per 1000	**248 per 1000** (134 to 459)				
	Moderate					
	181 per 1000	**248 per 1000** (134 to 460)				
Discontinuation due to adverse events (any switch versus continuing with the antidepressant) Number of participants discontinuing due to adverse events Follow-up: 6-12 weeks	**Study population**		**RR 1.74** (0.32 to 9.6)	546 (4 studies)	⊕⊖⊖⊖ **very low**[3,6,13,16]	
	19 per 1000	**33 per 1000** (6 to 183)				
	Moderate					
	20 per 1000	**35 per 1000** (6 to 192)				
Discontinuation due to adverse events (switch to SSRI versus continuing TCA/SNRI) Number of participants discontinuing due to	**Study population**		**RR 1.43** (0.38 to 5.47)	329 (2 studies)	⊕⊖⊖⊖ **very low**[3,15,16]	
	24 per 1000	**34 per 1000** (9 to 129)				
	Moderate					
	23 per 1000	**33 per 1000** (9 to 126)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to another antidepressant of a different class				
adverse events Follow-up: 8-12 weeks						
Discontinuation due to adverse events (switch to atypical AD/SNRI/TeCA [mianserin] versus continuing SSRI) Number of participants discontinuing due to adverse events Follow-up: 6-8 weeks	**Study population**		**RR 1.8** (0.01 to 222.73)	217 (2 studies)	⊕⊖⊖⊖ **very low**[3,16,17,18]	
	12 per 1000	**22 per 1000** (0 to 1000)				
	Moderate					
	11 per 1000	**20 per 1000** (0 to 1000)				

Notes:
[1] Risk of randomisation method is high risk or unclear, method of allocation concealment is unclear, intervention administration is non-blind, risk of detection bias is high or unclear, in studies contributing>50% to weighting in analysis
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] Funded by pharmaceutical company
[4] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration and outcome assessment
[5] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and unclear blinding of intervention administration and unclear blinding or non-blind outcome assessment, in studies contributing >50% to weighting in analysis
[6] I-squared>50%
[7] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[8] Unclear randomisation method and method of allocation concealment, unclear blinding of intervention administration, non-blind outcome assessment and unclear risk of attrition bias (drop-out>20% but difference between groups <20% and ITT analysis used)
[9] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[10] Unclear randomisation method and method of allocation concealment, unclear blinding of intervention administration, unclear blinding or non-blind outcome assessment, and unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[11] N<400
[12] 95% CI crosses both line of no effect and threshold for clinically important benefit (SMD -0.5)
[13] Unclear or high risk of bias associated with randomisation method, method of allocation concealment is unclear and unclear blinding of intervention administration
[14] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)
[15] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration
[16] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[17] Risk of randomisation method is high or unclear and unclear blinding of intervention administration
[18] I-squared>80%

Update 2017

Table 133: Study information table for trials included in the meta-analysis of switching to a non-antidepressant agent versus continuing with the antidepressant

	Switch to antipsychotic monotherapy versus continuing SSRI/TCA/SNRI	Switch to combined antipsychotic + SSRI versus continuing TCA/SNRI
Total no. of studies (N randomised)	3 (1588)	2 (983)
Study ID	Corya 2006[1] Shelton 2005[2] Thase 2007[3]	Corya 2006[1] Shelton 2005[2]
Country	16 countries[1] US and Canada[2,3]	16 countries[1] US and Canada[2]
Diagnostic status	DSM-IV MDD (single episode or recurrent), without psychotic features[1] DSM-IV unipolar nonpsychotic MDD, confirmed with the SCID[2] DSM-IV MDD (recurrent), without psychotic features, confirmed by the SCID-I[3]	DSM-IV MDD (single episode or recurrent), without psychotic features[1] DSM-IV unipolar nonpsychotic MDD, confirmed with the SCID[2]
Age range (mean)	Range NR (45.7) [1] Range NR (42.4) [2] 18-65 (44.4) [3]	Range NR (45.7) [1] Range NR (42.4) [2]
Sex (% female)	73[1] 68[2] 63[3]	73[1] 68[2]
Ethnicity (% BME)	10[1] 12[2] 14[3]	10[1] 12[2]
Mean age (SD) at first onset of depression	NR	NR
Mean months (SD) since onset of current episode	Median 26.6[1] Median: 11.8[2] 57.7 (80.9) [3]	Median 26.6[1] Median: 11.8[2]
No. (SD) of previous depressive episodes	Mean NR (51% >3 episodes) [1] NR[2,3]	Mean NR (51% >3 episodes) [1] NR[2]
Details of inadequate response/treatment resistance	TRD: Inadequate response to a selective serotonin reuptake inhibitor (SSRI) antidepressant after at least 6 weeks of therapy at a therapeutic dose (i.e., citalopram, 40 mg/day; fluoxetine, 40 mg/day; paroxetine, 40 mg/day; or sertraline, 150 mg/day) at entry into the trial and inadequate response (<30% improvement in MADRS total score) to an open-label, 7-week lead-in phase of venlafaxine (75–375 mg/day according to the investigator's clinical judgment)[1]	TRD: Inadequate response to a selective serotonin reuptake inhibitor (SSRI) antidepressant after at least 6 weeks of therapy at a therapeutic dose (i.e., citalopram, 40 mg/day; fluoxetine, 40 mg/day; paroxetine, 40 mg/day; or sertraline, 150 mg/day) at entry into the trial and inadequate response (<30% improvement in MADRS total score) to an open-label, 7-week lead-in phase of venlafaxine (75–375 mg/day according to the investigator's clinical judgment)[1]

Update 2017

	Switch to antipsychotic monotherapy versus continuing SSRI/TCA/SNRI	Switch to combined antipsychotic + SSRI versus continuing TCA/SNRI
	TRD: History of at least one failure to respond to SSRI after at least 4 weeks of therapy at a therapeutic dose (i.e. citalopram 40mg/day, fluoxetine 40mg/day, paroxetine 40mg/day, or sertraline 150mg/day), and failure to respond (<30% improvement on MADRS) to nortriptyline (25-175mg/day; mean modal dose 104.6mg/day) during a 7-week open-label treatment phase[2] TRD: Documented history of failure to achieve a satisfactory response (based on investigator's clinical judgement) to an antidepressant (except fluoxetine) after at least 6 weeks of therapy at a therapeutic dose (e.g. paroxetine 40mg/day, venlafaxine 150mg/day, bupropion 300mg/day, trazodone 450mg/day), and failure to respond (<25% decrease in HAMD) to an 8-week, open-label prospective fluoxetine (25-50mg/day) therapy lead-in[3]	TRD: History of at least one failure to respond to SSRI after at least 4 weeks of therapy at a therapeutic dose (i.e. citalopram 40mg/day, fluoxetine 40mg/day, paroxetine 40mg/day, or sertraline 150mg/day), and failure to respond (<30% improvement on MADRS) to nortriptyline (25-175mg/day; mean modal dose 104.6mg/day) during a 7-week open-label treatment phase[2]
Augmented/previous treatment	Previous treatment: Venlafaxine (75–375 mg/day). Mean number of prior psychotropic medications: 4.1[1] Previous treatment: Nortriptyline (25-175mg/day; mean modal dose 104.6mg/day)[2] Previous treatment: Fluoxetine (25-50mg/day)[3]	Previous treatment: Venlafaxine (75–375 mg/day). Mean number of prior psychotropic medications: 4.1[1] Previous treatment: Nortriptyline (25-175mg/day; mean modal dose 104.6mg/day)[2]
Baseline severity	MADRS 30 (More severe)[1,3] MADRS 28.5 (More severe)[2]	MADRS 30 (More severe)[1] MADRS 28.5 (More severe)[2]
Intervention details (mean dose)	Olanzapine 6 or 12mg/day (mean modal dose 7.9 mg/day)[1] Olanzapine 6-12mg/day (mean modal dose 8.3mg/day)[2] Olanzapine 6, 12 or 18mg/day (mean modal dose 8.7mg/day)[3]	Olanzapine 6 or 12 mg/day (mean modal dose 7.9 mg/day) + Fluoxetine 25 or 50mg/day (mean modal dose 37.5 mg/day)[1] Olanzapine: 6-12mg/day (mean modal dose 8.5mg/day) + Fluoxetine: 25-50mg/day (mean modal dose 35.6mg/day)[2]

Update 2017

	Switch to antipsychotic monotherapy versus continuing SSRI/TCA/SNRI	Switch to combined antipsychotic + SSRI versus continuing TCA/SNRI
Comparator details (mean dose, if applicable)	Venlafaxine 75-375mg/day (mean modal dose 275.4 mg/day) [1] Nortriptyline 25-175mg/day (mean modal dose 103.5mg/day) [2] Fluoxetine 50mg/day (mean modal dose 49.5mg/day) [3]	Venlafaxine 75-375mg/day (mean modal dose 275.4 mg/day) [1] Nortriptyline 25-175mg/day (mean modal dose 103.5mg/day) [2]
Treatment length (weeks)	12[1] 8[2,3]	12[1] 8[2]

Notes:
Abbreviations: mg=milligram, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression
[1]Corya 2006; [2]Shelton 2005; [3]Thase 2007
Note that Corya 2006[1] is a five-armed trial, Shelton 2005[2] is a four-armed trial and Thase 2007[3] is a three-armed trial and demographics reported here are for all arms combined

Table 134: Summary of findings table for switching to a non-antidepressant agent versus continuing with the antidepressant

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to non-antidepressant agent				
Remission (switch to antipsychotic monotherapy versus continuing SSRI/TCA/SNRI) Number of people scoring ≤8/10 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 0.79** (0.56 to 1.11)	729 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	179 per 1000	**142 per 1000** (100 to 199)				
	Moderate					
	177 per 1000	**140 per 1000** (99 to 196)				
Remission (switch to combined antipsychotic + SSRI versus continuing TCA/SNRI) Number of people scoring ≤8 on Montgomery Asberg Depression Rating Scale (MADRS)	**Study population**		**RR 1.17** (0.79 to 1.75)	502 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,4]	
	198 per 1000	**232 per 1000** (157 to 347)				
	Moderate					
	200 per 1000	**234 per 1000** (158 to 350)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to non-antidepressant agent				
Follow-up: 8-12 weeks						
Response (switch to antipsychotic monotherapy versus continuing SSRI/TCA/SNRI) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 0.69** (0.49 to 0.96)	729 (3 studies)	⊕⊖⊖⊖ **very low**[1,3,5]	
	334 per 1000	**231 per 1000** (164 to 321)				
	Moderate					
	309 per 1000	**213 per 1000** (151 to 297)				
Response (switch to combined antipsychotic + SSRI versus continuing TCA/SNRI) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 0.87** (0.68 to 1.12)	502 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	397 per 1000	**345 per 1000** (270 to 444)				
	Moderate					
	404 per 1000	**351 per 1000** (275 to 452)				
Depression symptomatology (switch to antipsychotic monotherapy versus continuing SSRI/TCA/SNRI) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 8-12 weeks		The mean depression symptomatology (switch to antipsychotic monotherapy versus continuing SSRI/TCA/SNRI) in the intervention groups was **0.22 standard deviations higher** (0.12 lower to 0.57 higher)		733 (3 studies)	⊕⊖⊖⊖ **very low**[1,3,6,7]	SMD 0.22 (-0.12 to 0.57)
Depression symptomatology (switch to		The mean depression symptomatology		516 (2 studies)	⊕⊕⊖⊖ **low**[1,3]	SMD -0.09 (-0.29 to 0.11)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to non-antidepressant agent				
combined antipsychotic + SSRI versus continuing TCA/SNRI) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 8-12 weeks		(switch to combined antipsychotic + SSRI versus continuing TCA/SNRI) in the intervention groups was **0.09 standard deviations lower** (0.29 lower to 0.11 higher)				
Discontinuation for any reason (switch to antipsychotic monotherapy versus continuing SSRI/TCA/SNRI) Number of participants discontinuing for any reason (including adverse events) Follow-up: 8-12 weeks	**Study population**		**RR 1.67** (1.26 to 2.23)	738 (3 studies)	⊕⊖⊖⊖ **very low**[3,5,8]	
	189 per 1000	**316 per 1000** (238 to 422)				
	Moderate					
	194 per 1000	**324 per 1000** (244 to 433)				
Discontinuation for any reason (switch to combined antipsychotic + SSRI versus continuing TCA/SNRI) Number of participants discontinuing for any reason (including adverse events) Follow-up: 8-12 weeks	**Study population**		**RR 1.22** (0.69 to 2.16)	516 (2 studies)	⊕⊖⊖⊖ **very low**[3,8,9]	
	181 per 1000	**221 per 1000** (125 to 391)				
	Moderate					
	186 per 1000	**227 per 1000** (128 to 402)				
Discontinuation due to adverse events (switch to antipsychotic monotherapy versus continuing	**Study population**		**RR 5.34** (2.57 to 11.09)	738 (3 studies)	⊕⊖⊖⊖ **very low**[3,5,8]	
	24 per 1000	**128 per 1000** (62 to 266)				
	Moderate					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Continuing with the antidepressant	Switch to non-antidepressant agent				
SSRI/TCA/SNRI) Number of participants discontinuing due to adverse events Follow-up: 8-12 weeks	**24 per 1000**	**128 per 1000** (62 to 266)				
Discontinuation due to adverse events (switch to combined antipsychotic + SSRI versus continuing TCA/SNRI) Number of participants discontinuing due to adverse events Follow-up: 8-12 weeks	**Study population**		**RR 3.48** (1.06 to 11.44)	516 (2 studies)	⊕⊖⊖⊖ **very low**[3,5,8]	
	24 per 1000	**82 per 1000** (25 to 270)				
	Moderate					
	23 per 1000	**80 per 1000** (24 to 263)				

Notes:
[1] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration and outcome assessment, and unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used) in studies contributing >50% to weighting in analysis
[2] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[3] Funding from pharmaceutical companies
[4] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[5] Events<300
[6] I-squared>50%
[7] 95% CI crosses both line of no effect and threshold for clinically important harm (SMD 0.5)
[8] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration
[9] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 135: Study information table for trials included in the meta-analysis of switching to another antidepressant or non-antidepressant agent versus augmenting with another antidepressant or non-antidepressant agent

	Switch to TeCA versus augmentation with TeCA (mianserin)	Switch to antipsychotic versus augmentation with antipsychotic	Switch to antipsychotic versus augmentation with lithium
Total no. of studies (N randomised)	1 (104)	2 (1293)	1 (688)
Study ID	Ferreri 2001	Bauer 2010/2013[1] Thase 2007[2]	Bauer 2010/2013

Update 2017

	Switch to TeCA versus augmentation with TeCA (mianserin)	Switch to antipsychotic versus augmentation with antipsychotic	Switch to antipsychotic versus augmentation with lithium
Country	France	Australia, Austria, Belgium, Bulgaria, Germany, Hungary, Italy, Portugal, Romania, Slovakia, Spain and the UK[1] US and Canada[2]	Australia, Austria, Belgium, Bulgaria, Germany, Hungary, Italy, Portugal, Romania, Slovakia, Spain and the UK
Diagnostic status	DSM-III-R MDD	DSM-IV diagnosis of MDD (single or recurrent episode), confirmed by the Mini International Neuropsychiatric Interview (MINI) [1] DSM-IV MDD (recurrent), without psychotic features, confirmed by the SCID-I[2]	DSM-IV diagnosis of MDD (single or recurrent episode), confirmed by the Mini International Neuropsychiatric Interview (MINI)
Age range (mean)	Range NR (46.6)	NR[1] 18-65 (44.4) [2]	NR
Sex (% female)	74	NR[1] 63[2]	NR
Ethnicity (% BME)	NR	NR[1] 14[2]	NR
Mean age (SD) at first onset of depression	NR	NR	NR
Mean months (SD) since onset of current episode	7.3 (8.4)	6 (3.8) [1] 57.7 (80.9) [2]	6 (3.8)
No. (SD) of previous depressive episodes	2.4 (2.2)	4.0 (6.0) [1] NR[2]	4.0 (6.0)
Details of inadequate response/treatment resistance	Inadequate response to previous fluoxetine treatment after at least 6 weeks of treatment with 20 mg/day	Patients were required to have Stage I or II TRD, 50% of participants fell into each category (defined as: Stage I-failure of ≥1 adequate trial of one major class of AD [citalopram, escitalopram, paroxetine, sertraline or venlafaxine]; Stage II-failure of adequate trials of two different classes of AD, the most recent of which must have been an AD listed for patients with Stage I TRD). An inadequate response was defined as not	Patients were required to have Stage I or II TRD, 50% of participants fell into each category (defined as: Stage I-failure of ≥1 adequate trial of one major class of AD [citalopram, escitalopram, paroxetine, sertraline or venlafaxine]; Stage II-failure of adequate trials of two different classes of AD, the most recent of which must have been an AD listed for patients with Stage I TRD). An inadequate response was defined as not

Update 2017

	Switch to TeCA versus augmentation with TeCA (mianserin)	Switch to antipsychotic versus augmentation with antipsychotic	Switch to antipsychotic versus augmentation with lithium
		achieving remission from depressive symptoms after receiving at least a minimum effective dose of an AD with ≥1 dose increase for ≥28 days prior to the study[1] TRD: Documented history of failure to achieve a satisfactory response (based on investigator's clinical judgement) to an antidepressant (except fluoxetine) after at least 6 weeks of therapy at a therapeutic dose (e.g. paroxetine 40mg/day, venlafaxine 150mg/day, bupropion 300mg/day, trazodone 450mg/day), and failure to respond (<25% decrease in HAMD) to an 8-week, open-label prospective fluoxetine (25-50mg/day) therapy lead-in[2]	achieving remission from depressive symptoms after receiving at least a minimum effective dose of an AD with ≥1 dose increase for ≥28 days prior to the study
Augmented/previous treatment	Augmented/previous antidepressant: Fluoxetine (20mg/day)	Augmented/previous antidepressant: 66% SSRI; 36% venlafaxine; 8% other AD[1] Augmented/previous antidepressant: Fluoxetine[2]	Augmented/previous antidepressant: 66% SSRI; 36% venlafaxine; 8% other AD
Baseline severity	HAMD 27.2 (More severe)	MADRS 33.3 (More severe) [1] MADRS 30 (More severe) [2]	MADRS 33.3 (More severe)
Intervention details (mean dose)	Mianserin 60mg/day	Quetiapine extended-release (XR) 200-300mg/day (titrated upwards from 50mg/day to 300mg/day in first week and titrated downwards if necessary; mean dose 238mg/day [SD=60]) [1] Olanzapine 6, 12 or 18mg/day (mean	Quetiapine extended-release (XR) 200-300mg/day (titrated upwards from 50mg/day to 300mg/day in first week and titrated downwards if necessary; mean dose 238mg/day [SD=60])

Update 2017

	Switch to TeCA versus augmentation with TeCA (mianserin)	Switch to antipsychotic versus augmentation with antipsychotic	Switch to antipsychotic versus augmentation with lithium
		modal dose 8.7mg/day)[2]	
Comparator details (mean dose, if applicable)	Mianserin 60mg/day + Fluoxetine: 20mg/day	Quetiapine extended-release (XR) 200-300mg/day (titrated upwards from 50mg/day to 300mg/day in first week and titrated downwards if necessary; mean dose 242mg/day [SD=54]) + usual AD (SSRI/venlafaxine)[1] Olanzapine 6, 12 or 18mg/day (mean modal dose 8.6mg/day) + fluoxetine 50mg/day (mean modal dose 48.8mg/day)[2]	Lithium 450-900mg/day (target plasma level: 0.6–1.2mmol/L; mean dose 882mg/day [SD=212]) + usual AD (SSRI/venlafaxine)
Treatment length (weeks)	6	6[1] 8[2]	6

Notes:
Abbreviations: mg=milligram, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression
[1]Bauer 2010/2013; [2]Thase 2007
Note that Bauer 2010/2013, Ferreri 2001 and Thase 2007 are three-armed trials and demographics reported here are for the three arms combined

Update 2017

Table 136: Summary of findings table for switching to another antidepressant or non-antidepressant agent versus augmenting with another antidepressant or non-antidepressant agent

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmentation with another antidepressant/non-antidepressant agent	Switch to another antidepressant/non-antidepressant agent				
Remission (switch to TeCA versus augmentation with TeCA [mianserin])	**Study population**		**RR 0.83** (0.46 to 1.51)	65 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	438 per 1000	**363 per 1000** (201 to 661)				
	Moderate					

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmentation with another antidepressant/non-antidepressant agent	**Switch to another antidepressant/non-antidepressant agent**				
Number of people scoring ≤8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 6 weeks	**438 per 1000**	**364 per 1000** (201 to 661)				
Remission (switch to antipsychotic versus augmentation with antipsychotic) Number of people scoring ≤10/<10 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 6-8 weeks	**Study population**		**RR 0.65** (0.48 to 0.88)	849 (2 studies)	⊕⊖⊖⊖ **very low**[4,5,6]	
	297 per 1000	**193 per 1000** (143 to 262)				
	Moderate					
	296 per 1000	**192 per 1000** (142 to 260)				
Remission (switch to antipsychotic versus augmentation with lithium) Number of people scoring <10 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 6 weeks	**Study population**		**RR 0.87** (0.63 to 1.19)	446 (1 study)	⊕⊖⊖⊖ **very low**[6,7,8]	
	271 per 1000	**236 per 1000** (171 to 323)				
	Moderate					
	272 per 1000	**237 per 1000** (171 to 324)				
Response (switch to TeCA versus augmentation with TeCA [mianserin])	**Study population**		**RR 0.78** (0.5 to 1.21)	65 (1 study)	⊕⊖⊖⊖ **very low**[1,3,8]	
	625 per 1000	**488 per 1000** (312 to 756)				
	Moderate					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmentation with another antidepressant/non-antidepressant agent	**Switch to another antidepressant/non-antidepressant agent**				
Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 6 weeks	**625 per 1000**	**488 per 1000** (312 to 756)				
Response (switch to antipsychotic versus augmentation with antipsychotic) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 6-8 weeks	**Study population**		**RR 0.8** (0.53 to 1.2)	849 (2 studies)	⊕⊖⊖⊖ **very low**[4,6,8,9]	
	468 per 1000	**375 per 1000** (248 to 562)				
	Moderate					
	464 per 1000	**371 per 1000** (246 to 557)				
Response (switch to antipsychotic versus augmentation with lithium) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 6 weeks	**Study population**		**RR 1** (0.83 to 1.2)	446 (1 study)	⊕⊖⊖⊖ **very low**[5,6,7]	
	507 per 1000	**507 per 1000** (421 to 608)				
	Moderate					
	507 per 1000	**507 per 1000** (421 to 608)				
	Study population					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmentation with another antidepressant/non-antidepressant agent	**Switch to another antidepressant/non-antidepressant agent**				
Response (switch to TeCA versus augmentation with TeCA [mianserin]) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 6 weeks	**719 per 1000**	**640 per 1000** (453 to 891)				
	Moderate					
	719 per 1000	**640 per 1000** (453 to 892)	**RR 0.89** (0.63 to 1.24)	65 (1 study)	⊕⊖⊖⊖ **very low**[1,3,8]	
Response (switch to antipsychotic versus augmentation with antipsychotic) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 6 weeks	**Study population**		**RR 0.92** (0.81 to 1.06)	454 (1 study)	⊕⊖⊖⊖ **very low**[5,6,7]	
	668 per 1000	**615 per 1000** (541 to 708)				
	Moderate					
	668 per 1000	**615 per 1000** (541 to 708)				
Response (switch to antipsychotic versus augmentation with lithium) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 6 weeks	**Study population**		**RR 1.03** (0.88 to 1.19)	446 (1 study)	⊕⊖⊖⊖ **very low**[5,6,7]	
	602 per 1000	**620 per 1000** (530 to 716)				
	Moderate					
	602 per 1000	**620 per 1000** (530 to 716)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmentation with another antidepressant/non-antidepressant agent	**Switch to another antidepressant/non-antidepressant agent**				
Depression symptomatology (any switch versus any augmentation) Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 6-8 weeks		The mean depression symptomatology (any switch versus any augmentation) in the intervention groups was **0.39 standard deviations higher** (0.2 to 0.57 higher)		460 (2 studies)	⊕⊕⊖⊖ **low**[3,10]	SMD 0.39 (0.2 to 0.57)
Depression symptomatology (switch to TeCA versus augmentation with TeCA [mianserin]) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 6 weeks		The mean depression symptomatology (switch to TeCA versus augmentation with TeCA [mianserin]) in the intervention groups was **0.41 standard deviations higher** (0.08 lower to 0.91 higher)		65 (1 study)	⊕⊖⊖⊖ **very low**[1,3,11]	SMD 0.41 (-0.08 to 0.91)
Depression symptomatology (switch to antipsychotic versus augmentation with antipsychotic) Montgomery Asberg Depression Rating Scale (MADRS; change score)		The mean depression symptomatology (switch to antipsychotic versus augmentation with antipsychotic) in the intervention groups was **0.38 standard deviations higher** (0.18 to 0.58 higher)		395 (1 study)	⊕⊖⊖⊖ **very low**[3,12,13]	SMD 0.38 (0.18 to 0.58)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmentation with another antidepressant/non-antidepressant agent	**Switch to another antidepressant/non-antidepressant agent**				
Follow-up: mean 8 weeks						
Discontinuation for any reason (switch to TeCA versus augmentation with TeCA [mianserin]) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 6 weeks	**Study population**		**RR 1.88** (0.8 to 4.42)	66 (1 study)	⊕⊖⊖⊖ **very low**[3,14,15]	
	188 per 1000	**352 per 1000** (150 to 829)				
	Moderate					
	188 per 1000	**353 per 1000** (150 to 831)				
Discontinuation for any reason (switch to antipsychotic versus augmentation with antipsychotic) Number of participants discontinuing for any reason (including adverse events) Follow-up: 6-8 weeks	**Study population**		**RR 1.4** (1.11 to 1.78)	858 (2 studies)	⊕⊖⊖⊖ **very low**[4,5,6]	
	202 per 1000	**283 per 1000** (224 to 359)				
	Moderate					
	206 per 1000	**288 per 1000** (229 to 367)				
Discontinuation for any reason (switch to antipsychotic versus augmentation with lithium) Number of participants discontinuing for any reason (including adverse	**Study population**		**RR 1.05** (0.73 to 1.49)	457 (1 study)	⊕⊖⊖⊖ **very low**[6,7,16]	
	205 per 1000	**216 per 1000** (150 to 306)				
	Moderate					
	205 per 1000	**215 per 1000** (150 to 305)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmentation with another antidepressant/non-antidepressant agent	**Switch to another antidepressant/non-antidepressant agent**				
events) Follow-up: mean 6 weeks						
Discontinuation due to adverse events (switch to TeCA versus augmentation with TeCA [mianserin]) Number of participants discontinuing due to adverse events Follow-up: mean 6 weeks	**Study population** **62 per 1000** **Moderate** **63 per 1000**	 **235 per 1000** (54 to 1000) **237 per 1000** (54 to 1000)	**RR 3.76** (0.86 to 16.41)	66 (1 study)	⊕⊖⊖⊖ **very low**[3,14,15]	
Discontinuation due to adverse events (switch to antipsychotic versus augmentation with antipsychotic) Number of participants discontinuing due to adverse events Follow-up: 6-8 weeks	**Study population** **116 per 1000** **Moderate** **117 per 1000**	 **140 per 1000** (99 to 200) **142 per 1000** (99 to 201)	**RR 1.21** (0.85 to 1.72)	858 (2 studies)	⊕⊖⊖⊖ **very low**[4,6,15]	
Discontinuation due to adverse events (switch to antipsychotic versus augmentation with lithium) Number of participants discontinuing due to adverse events Follow-up: mean 6 weeks	**Study population** **79 per 1000** **Moderate** **79 per 1000**	 **123 per 1000** (70 to 215) **123 per 1000** (70 to 216)	**RR 1.56** (0.89 to 2.74)	457 (1 study)	⊕⊖⊖⊖ **very low**[6,7,15]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Augmentation with another antidepressant/non-antidepressant agent	**Switch to another antidepressant/non-antidepressant agent**				

Notes:
[1] Unclear randomisation method and method of allocation concealment and unclear blinding of intervention administration. Risk of attrition bias was also unclear (drop-out>20% but difference between groups<20% and ITT analysis used). Outcome assessment was non-blind
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] Funding from pharmaceutical companies
[4] Unclear method of allocation concealment and unclear blinding of, or non-blind, intervention administrator(s)
[5] Events<300
[6] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical companies
[7] Unclear method of allocation concealment and non-blind intervention administrator(s)
[8] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[9] I-squared>80%
[10] Unclear randomisation method and method of allocation concealment, unclear blinding of intervention administrator(s), unclear blinding of (or non-blind) outcome assessment, and unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[11] 95% CI crosses both line of no effect and threshold for clinically important harm (SMD 0.5)
[12] Unclear randomisation method and method of allocation concealment, unclear blinding of intervention administrator(s) and outcome assessment, and unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[13] N<400
[14] Unclear randomisation method and method of allocation concealment and unclear blinding of intervention administration
[15] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)
[16] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Update 2017

Table 137: Study information table for trials included in the meta-analysis of switching to another antidepressant of the same class versus switching to another antidepressant of a different class

	Switch to another SSRI versus switch to SNRI	Switch to another SSRI versus switch to an atypical AD
Total no. of studies (N randomised)	2 (1133)	1 (727)
Study ID	Lenox-Smith 2008[1] Rush 2006[2]	Rush 2006
Country	Europe and Australia[1] US[2]	US
Diagnostic status	DSM-IV MDD[1] DSM-IV nonpsychotic MDD[2]	DSM-IV nonpsychotic MDD
Age range (mean)	Range NR (42.5) [1] Range NR (41.8) [2]	Range NR (41.8)

	Switch to another SSRI versus switch to SNRI	Switch to another SSRI versus switch to an atypical AD
Sex (% female)	67[1] 59[2]	59
Ethnicity (% BME)	NR[1] 24[2]	24
Mean age (SD) at first onset of depression	NR[1] 25.0 (14.0) [2]	25.0 (14.0)
Mean months (SD) since onset of current episode	NR[1] 29.6 (65.9). 27% chronic MDD (≥2 years) [2]	29.6 (65.9). 27% chronic MDD (≥2 years)
No. (SD) of previous depressive episodes	NR[1] 7.0 (12.8) [2]	7.0 (12.8)
Details of inadequate response/treatment resistance	Inadequate response following 8 weeks of monotherapy with an adequate dosing regimen of an SSRI other than citalopram[1] Inadequate response (not achieved remission or who were intolerant [56%]) to an initial prospective treatment with citalopram[2]	Inadequate response (not achieved remission or who were intolerant [56%]) to an initial prospective treatment with citalopram
Augmented/previous treatment	Previous treatment: SSRI (not citalopram) [1] Previous treatment: Citalopram[2]	Previous treatment: Citalopram
Baseline severity	MADRS 30.9 (More severe) [1] HAMD 18.9 (Less severe) [2]	HAMD 18.9 (Less severe)
Intervention details (mean dose)	Citalopram 20-60mg/day (final mean dose 51 mg/day) [1] Sertraline 50-200mg/day (mean final dose 135.5mg [SD=57.4]) [2]	Sertraline 50-200mg/day (mean final dose 135.5mg [SD=57.4])
Comparator details (mean dose, if applicable)	Venlafaxine extended release 75-300mg/day (final mean dose 191 mg/day) [1] Venlafaxine extended release 37.5-375mg/day (mean final dose 193.6mg [SD=106.2]) [2]	Venlafaxine extended release 37.5-375mg/day (mean final dose 193.6mg [SD=106.2])
Treatment length (weeks)	12[1] 14[2]	14
Notes: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression [1]Lenox-Smith 2008; [2]Rush 2006 Note that Rush 2006 is a three-armed trial and demographics reported here are for all arms combined		

Update 2017

Table 138: Summary of findings table for switching to another antidepressant of the same class versus switching to another antidepressant of a different class

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	Switch to another antidepressant of the same class versus switch to another antidepressant of a different class				
Remission (switch to another SSRI versus switch to SNRI) Number of people scoring ≤4/7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 12-14 weeks	**Study population**		**RR 0.61** (0.45 to 0.83)	884 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	277 per 1000	**169 per 1000** (125 to 230)				
	Moderate					
	281 per 1000	**171 per 1000** (126 to 233)				
Remission (switch to another SSRI versus switch to an atypical AD) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 14 weeks	**Study population**		**RR 0.83** (0.57 to 1.19)	477 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	213 per 1000	**177 per 1000** (122 to 254)				
	Moderate					
	213 per 1000	**177 per 1000** (121 to 253)				
Response (switch to another SSRI versus switch to SNRI) Number of people showing ≥50% improvement on Quick Inventory of Depressive Symptomatology (QIDS) Follow-up: mean 14 weeks	**Study population**		**RR 0.95** (0.71 to 1.26)	488 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	
	280 per 1000	**266 per 1000** (199 to 353)				
	Moderate					
	365 per 1000	**347 per 1000** (259 to 460)				
Response (switch to another SSRI versus switch to an atypical AD) Number of people showing ≥50%	**Study population**		**RR 1.02** (0.76 to 1.38)	477 (1 study)	⊕⊖⊖⊖ **very low**[1,3,6]	
	259 per 1000	**265 per 1000** (197 to 358)				
	Moderate					

Update 2017

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	Switch to another antidepressant of the same class versus switch to another antidepressant of a different class				
improvement on Quick Inventory of Depressive Symptomatology (QIDS) Follow-up: mean 14 weeks	**259 per 1000**	**264 per 1000** (197 to 357)				
Depression symptomatology (switch to another SSRI versus switch to SNRI) Quick Inventory of Depressive Symptomatology (QIDS; change score) Follow-up: mean 14 weeks		The mean depression symptomatology (switch to another SSRI versus switch to SNRI) in the intervention groups was **0.08 standard deviations lower** (0.26 lower to 0.09 higher)		488 (1 study)	⊕⊕⊖⊖ **low**[1,3]	SMD -0.08 (-0.26 to 0.09)
Depression symptomatology (switch to another SSRI versus switch to an atypical AD) Quick Inventory of Depressive Symptomatology (QIDS; change score) Follow-up: mean 14 weeks		The mean depression symptomatology (switch to another SSRI versus switch to an atypical ad) in the intervention groups was **0.12 standard deviations lower** (0.3 lower to 0.06 higher)		477 (1 study)	⊕⊕⊖⊖ **low**[1,3]	SMD -0.12 (-0.3 to 0.06)
Discontinuation for any reason (switch to another SSRI versus switch to SNRI) Number of participants discontinuing for any reason including adverse events Follow-up: mean 12 weeks	**Study population**		**RR 0.85** (0.59 to 1.22)	406 (1 study)	⊕⊖⊖⊖ **very low**[1,3,7]	
	245 per 1000	**208 per 1000** (145 to 299)				
	Moderate					
	245 per 1000	**208 per 1000** (145 to 299)				
	Study population					

Update 2017

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	Switch to another antidepressant of the same class versus switch to another antidepressant of a different class				
Discontinuation due to adverse events (switch to another SSRI versus switch to SNRI) Number of participants discontinuing due to adverse events Follow-up: 12-14 weeks	**143 per 1000**	**141 per 1000** (103 to 193)	**RR 0.99** (0.72 to 1.35)	891 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,8]	
	Moderate					
	134 per 1000	**133 per 1000** (96 to 181)				
Discontinuation due to adverse events (switch to another SSRI versus switch to an atypical AD) Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Study population**		**RR 0.77** (0.56 to 1.07)	477 (1 study)	⊕⊖⊖⊖ **very low**[1,3,7]	
	272 per 1000	**209 per 1000** (152 to 291)				
	Moderate					
	272 per 1000	**209 per 1000** (152 to 291)				

Notes:
[1] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administrator(s)
[2] Events<300
[3] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical companies
[4] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[5] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and for clinically important benefit (RR 1.25)
[6] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[7] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 0.75)
[8] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Update 2017

Table 139: Study information table for trials included in the meta-analysis of switching to another antidepressant or non-antidepressant agent – head-to-head comparisons (part 1)

	Switch to SSRI versus switch to non-SSRI AD	Switch to SSRI versus switch to antipsychotic	Switch to SNRI versus switch to atypical antidepressant
Total no. of studies (N randomised)	4 (1445)	2 (983)	2 (1102)
Study ID	Lenox-Smith 2008[1] Poirier 1999[2] Rush 2006[3] Souery 2011a[4]	Corya 2006[5] Shelton 2005[6]	Fang 2010/2011[7] Rush 2006[3]
Country	Europe and Australia[1] France[2] US[3] Austria, Belgium, France and Israel[4]	16 countries[5] US and Canada[6]	China[7] US[3]
Diagnostic status	DSM-IV MDD[1] DSM-III-R MDD[2] DSM-IV nonpsychotic MDD[3] DSM-IV major depressive episode[4]	DSM-IV MDD (single episode or recurrent), without psychotic features[5] DSM-IV unipolar nonpsychotic MDD, confirmed with the SCID[6]	DSM-IV MDD[7] DSM-IV nonpsychotic MDD[3]
Age range (mean)	Range NR (42.5)[1] 21-61 (43.3)[2] Range NR (41.8)[3] Range NR (51.4)[4]	Range NR (45.7)[5] Range NR (42.4)[6]	NR[7] Range NR (41.8)[3]
Sex (% female)	67[1] 72[2,4] 59[3]	73[5] 68[6]	NR[7] 59[3]
Ethnicity (% BME)	NR[1,2] 24[3] 5[4]	10[5] 12[6]	NR[7] 24[3]
Mean age (SD) at first onset of depression	NR[1,2] 25.0 (14.0)[3] 38.8 (16.2)[4]	NR	NR[7] 25.0 (14.0)[3]
Mean months (SD) since onset of current episode	NR[1,4] 0.4 (0.2)[2] 29.6 (65.9). 27% chronic MDD (≥2 years)[3]	Median 26.6[5] Median: 11.8[6]	NR[7] 29.6 (65.9). 27% chronic MDD (≥2 years)[3]
No. (SD) of previous depressive episodes	NR[1,2] 7.0 (12.8)[3] 3.6 (4.2)[4]	Mean NR (51% >3 episodes)[5] NR[6]	NR[7] 7.0 (12.8)[3]
Details of inadequate response/treatment resistance	Inadequate response following 8 weeks of monotherapy with an adequate dosing regimen of an SSRI other than citalopram[1]	TRD: Inadequate response to a selective serotonin reuptake inhibitor (SSRI) antidepressant after at least 6 weeks of	TRD: Inadequate response to to 2 or more adequate treatments from different classes of antidepressants in the

Update 2017

	Switch to SSRI versus switch to non-SSRI AD	Switch to SSRI versus switch to antipsychotic	Switch to SNRI versus switch to atypical antidepressant
	TRD: History of resistance to 2 previous successive antidepressant treatments for the current episode (except venlafaxine or paroxetine). The first treatment had to have been for at least 4 weeks at an effective dose. The second treatment had to have been prescribed by the investigator at an effective dose (equivalent to 100-150mg of clomipramine as judged by the investigator) for at least 4 weeks before the first day of the study, or for at least 2 weeks if a safety problem caused the discontinuation. Participants were to be no more than 'minimally improved' (CGI improvement score of 3) with their second treatment[2] Inadequate response (not achieved remission or who were intolerant [56%]) to an initial prospective treatment with citalopram[3] Inadequate response to treatment with at least one antidepressant given at an adequate dose for at least 4 weeks, except citalopram and desipramine[4]	therapy at a therapeutic dose (i.e., citalopram, 40 mg/day; fluoxetine, 40 mg/day; paroxetine, 40 mg/day; or sertraline, 150 mg/day) at entry into the trial and inadequate response (<30% improvement in MADRS total score) to an open-label, 7-week lead-in phase of venlafaxine (75–375 mg/day according to the investigator's clinical judgment)[5] TRD: History of at least one failure to respond to SSRI after at least 4 weeks of therapy at a therapeutic dose (i.e. citalopram 40mg/day, fluoxetine 40mg/day, paroxetine 40mg/day, or sertraline 150mg/day), and failure to respond (<30% improvement on MADRS) to nortriptyline (25-175mg/day; mean modal dose 104.6mg/day) during a 7-week open-label treatment phase[6]	current depressive episode (adequate dosages of antidepressants with at least 3-month duration) determined through medical records and/or prospective treatment[7] Inadequate response (not achieved remission or who were intolerant [56%]) to an initial prospective treatment with citalopram[3]
Augmented/previous treatment	Previous treatment: SSRI (not citalopram) [1] Previous treatment: 71% had used a TCA to treat current episode, while an SSRI had been used by 65%[2]	Previous treatment: Venlafaxine (75–375 mg/day). Mean number of prior psychotropic medications: 4.1[5] Previous treatment: Nortriptyline (25-	Previous treatment Paroxetine[7] Previous treatment: Citalopram[3]

Update 2017

	Switch to SSRI versus switch to non-SSRI AD	Switch to SSRI versus switch to antipsychotic	Switch to SNRI versus switch to atypical antidepressant
	Previous treatment: Citalopram[3] Previous treatment of current episode: 34% SSRIs; 21% TCAs; 15% SNRIs; 8% trazodone/nefazodone; 6% NASSAs; 6% NRIs; 2% MAOIs; 1% SSREs[4]	175mg/day; mean modal dose 104.6mg/day)[6]	
Baseline severity	MADRS 30.9 (More severe)[1] HAMD 24.6 (More severe)[2] HAMD 18.9 (Less severe)[3] MADRS 31.5 (More severe)[4]	MADRS 30 (More severe)[5] MADRS 28.5 (More severe)[6]	NR[7] HAMD 18.9 (Less severe)[3]
Intervention details (mean dose)	Citalopram 20-60mg/day (final mean dose 51 mg/day)[1] Paroxetine 20-40mg/day (mean dose 36.3 mg/day [SD=4.9])[2] Sertraline 50-200mg/day (mean final dose 135.5mg [SD=57.4])[3] Citalopram minimum dose of 40mg/day (mean final dose 43.06mg/day)[4]	Fluoxetine 25 or 50mg/day (mean modal dose 37.5 mg/day)[5] Fluoxetine 25-50mg/day (mean modal dose 35.8mg/day)[6]	Venlafaxine-XR 225 mg/day[7] Venlafaxine extended release 37.5-375mg/day (mean final dose 193.6mg [SD=106.2])[3]
Comparator details (mean dose, if applicable)	Venlafaxine extended release 75-300mg/day (final mean dose 191 mg/day)[1] Venlafaxine 65-300mg/day (mean dose 269.0 mg/day [SD=46.7])[2] Bupropion Sustained Release 150-400mg/day (mean final dose 282.7mg [SD=104.4]) or Venlafaxine extended release 37.5-375mg/day (mean final dose 193.6mg [SD=106.2])[3] Desipramine minimum dose 150mg/day	Olanzapine 6 or 12mg/day (mean modal dose 7.9 mg/day)[5] Olanzapine 6-12mg/day (mean modal dose 8.3mg/day)[6]	Mirtazapine 45mg/day[7] Bupropion Sustained Release 150-400mg/day (mean final dose 282.7mg [SD=104.4])[3]

Update 2017

	Switch to SSRI versus switch to non-SSRI AD	Switch to SSRI versus switch to antipsychotic	Switch to SNRI versus switch to atypical antidepressant
	(mean final dose 169.61mg/day) [4]		
Treatment length (weeks)	12[1] 4[2,4] 14[3]	12[5] 8[6]	8[7] 14[3]
Notes: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression [1]Lenox-Smith 2008; [2]Poirier 1999; [3]Rush 2006; [4]Souery 2011a; [5]Corya 2006; [6]Shelton 2005 Note that Corya 2006[5] is a five-armed trial, Fang 2010/2011[7] is an eight-armed trial, Rush 2006[3] is a three-armed trial and Shelton 2005[6] is a four-armed trial and demographics reported here are for all arms combined			

Table 140: Study information table for trials included in the meta-analysis of switching to another antidepressant or non-antidepressant agent – head-to-head comparisons (part 2)

	Switch to SSRI + antipsychotic versus switch to antipsychotic-only	Switch to SSRI + antipsychotic versus switch to SSRI-only
Total no. of studies (N randomised)	2 (983)	
Study ID	Corya 2006[1] Shelton 2005[2]	
Country	16 countries[1] US and Canada[2]	
Diagnostic status	DSM-IV MDD (single episode or recurrent), without psychotic features[1] DSM-IV unipolar nonpsychotic MDD, confirmed with the SCID[2]	
Age range (mean)	Range NR (45.7) [1] Range NR (42.4) [2]	
Sex (% female)	73[1] 68[2]	
Ethnicity (% BME)	10[1] 12[2]	
Mean age (SD) at first onset of depression	NR	
Mean months (SD) since onset of current episode	Median 26.6[1] Median: 11.8[2]	
No. (SD) of previous depressive episodes	Mean NR (51% >3 episodes) [1] NR[2]	
Details of inadequate response/treatment resistance	TRD: Inadequate response to a selective serotonin reuptake inhibitor (SSRI) antidepressant after at least 6 weeks of therapy at a therapeutic dose (i.e., citalopram, 40 mg/day; fluoxetine, 40 mg/day; paroxetine, 40 mg/day; or sertraline, 150 mg/day) at entry into the trial and inadequate response (<30% improvement in MADRS total score) to an open-label, 7-week lead-in phase of venlafaxine (75–375 mg/day according to the investigator's clinical judgment)[1]	

Update 2017

	Switch to SSRI + antipsychotic versus switch to antipsychotic-only	Switch to SSRI + antipsychotic versus switch to SSRI-only
	TRD: History of at least one failure to respond to SSRI after at least 4 weeks of therapy at a therapeutic dose (i.e. citalopram 40mg/day, fluoxetine 40mg/day, paroxetine 40mg/day, or sertraline 150mg/day), and failure to respond (<30% improvement on MADRS) to nortriptyline (25-175mg/day; mean modal dose 104.6mg/day) during a 7-week open-label treatment phase[2]	
Augmented/previous treatment	Previous treatment: Venlafaxine (75–375 mg/day). Mean number of prior psychotropic medications: 4.1[1] Previous treatment: Nortriptyline (25-175mg/day; mean modal dose 104.6mg/day)[2]	
Baseline severity	MADRS 30 (More severe)[1] MADRS 28.5 (More severe)[2]	
Intervention details (mean dose)	Fluoxetine 25 or 50mg/day (mean modal dose 37.5 mg/day) + Olanzapine: 6 or 12 mg/day (mean modal dose 7.9 mg/day)[1] Fluoxetine 25-50mg/day (mean modal dose 35.6mg/day) + Olanzapine: 6-12mg/day (mean modal dose 8.5mg/day)[2]	
Comparator details (mean dose, if applicable)	Olanzapine 6 or 12mg/day (mean modal dose 7.9 mg/day)[1] Olanzapine 6-12mg/day (mean modal dose 8.3mg/day)[2]	Fluoxetine 25 or 50mg/day (mean modal dose 37.5 mg/day)[1] Fluoxetine 25-50mg/day (mean modal dose 35.8mg/day)[2]
Treatment length (weeks)	12[1] 8[2]	
Notes: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression [1]Corya 2006; [2]Shelton 2005 Note that Corya 2006[1] is a five-armed trial and Shelton 2005[2] is a four-armed trial and demographics reported here are for all arms combined		

Table 141: Summary of findings table for switching to another antidepressant or non-antidepressant agent – head-to-head comparisons

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 Switch to another antidepressant/non-antidepressant agent (head-to-head)				
Remission (switch to SSRI versus switch to non-SSRI AD) Number of people scoring ≤4/≤7/<10 on Hamilton Rating Scale for Depression (HAM-D)	**Study population** **268 per 1000**	**166 per 1000** (134 to 206)	**RR 0.62** (0.5 to 0.77)	1397 (4 studies)	⊕⊕⊖⊖ **low**[1,2]	
	Moderate **314 per 1000**	**195 per 1000** (157 to 242)				

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	Switch to another antidepressant/non-antidepressant agent (head-to-head)				
Follow-up: 4-14 weeks						
Remission (switch to SSRI versus switch to antipsychotic) Number of people scoring ≤8 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 1.1** (0.68 to 1.8)	401 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	133 per 1000	**146 per 1000** (90 to 239)				
	Moderate					
	134 per 1000	**147 per 1000** (91 to 241)				
Remission (switch to SNRI versus switch to atypical antidepressant) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 8-14 weeks	**Study population**		**RR 1.16** (0.89 to 1.52)	594 (2 studies)	⊕⊖⊖⊖ **very low**[2,4,5]	
	241 per 1000	**280 per 1000** (215 to 367)				
	Moderate					
	289 per 1000	**335 per 1000** (257 to 439)				
Remission (switch to SSRI + antipsychotic versus switch to antipsychotic-only) Number of people scoring ≤8 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 1.63** (0.97 to 2.76)	579 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,5]	
	133 per 1000	**217 per 1000** (129 to 367)				
	Moderate					
	134 per 1000	**218 per 1000** (130 to 370)				
Remission (switch to SSRI + antipsychotic versus switch to SSRI-only) Number of people scoring ≤8 on Montgomery Asberg Depression Rating Scale (MADRS)	**Study population**		**RR 1.45** (0.97 to 2.17)	574 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,5]	
	146 per 1000	**212 per 1000** (142 to 318)				
	Moderate					
	156 per 1000	**226 per 1000** (151 to 339)				

Update 2017

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	Switch to another antidepressant/non-antidepressant agent (head-to-head)				
Follow-up: 8-12 weeks						
Response (switch to SSRI versus switch to non-SSRI AD) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Quick Inventory of Depressive Symptomatology (QIDS)/≥50% improvement on HAM-D AND much/very much improved on CGI-I (score 1-2) Follow-up: 4-14 weeks	**Study population** **318 per 1000** **Moderate** **450 per 1000**	**290 per 1000** (235 to 356) **410 per 1000** (333 to 504)	**RR 0.91** (0.74 to 1.12)	1001 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,6]	
Response (switch to SSRI versus switch to antipsychotic) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population** **212 per 1000** **Moderate** **224 per 1000**	**303 per 1000** (216 to 426) **320 per 1000** (228 to 450)	**RR 1.43** (1.02 to 2.01)	401 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,7]	
Response (switch to SNRI versus switch to atypical antidepressant) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Quick Inventory of Depressive Symptomatology (QIDS)	**Study population** **320 per 1000** **Moderate** **421 per 1000**	**349 per 1000** (281 to 432) **459 per 1000** (370 to 568)	**RR 1.09** (0.88 to 1.35)	594 (2 studies)	⊕⊖⊖⊖ **very low**[2,4,5]	

Update 2017

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	Switch to another antidepressant/non-antidepressant agent (head-to-head)				
Follow-up: 8-14 weeks						
Response (switch to SSRI + antipsychotic versus switch to antipsychotic-only) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population** **212 per 1000**	**326 per 1000** (239 to 445)	**RR 1.54** (1.13 to 2.1)	579 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,7]	
	Moderate **224 per 1000**	**345 per 1000** (253 to 470)				
Response (switch to SSRI + antipsychotic versus switch to SSRI-only) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population** **303 per 1000**	**330 per 1000** (248 to 445)	**RR 1.09** (0.82 to 1.47)	574 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,5]	
	Moderate **314 per 1000**	**342 per 1000** (257 to 462)				
Response (switch to SSRI versus switch to SNRI) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 4 weeks	**Study population** **635 per 1000**	**654 per 1000** (495 to 869)	**RR 1.03** (0.78 to 1.37)	107 (1 study)	⊕⊖⊖⊖ **very low**[1,2,5]	
	Moderate **635 per 1000**	**654 per 1000** (495 to 870)				
Depression symptomatology (switch to SSRI versus switch to non-SSRI AD) Hamilton Rating Scale for Depression (HAM-D; change		The mean depression symptomatology (switch to SSRI versus switch to non-SSRI ad) in the intervention groups was **0.08 standard deviations higher**		986 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,8]	SMD 0.08 (-0.18 to 0.34)

Update 2017

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	Switch to another antidepressant/non-antidepressant agent (head-to-head)				
score)/Quick Inventory of Depressive Symptomatology (QIDS; change score) Follow-up: 4-14 weeks		(0.18 lower to 0.34 higher)				
Depression symptomatology (switch to SSRI versus switch to antipsychotic) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 8-12 weeks		The mean depression symptomatology (switch to SSRI versus switch to antipsychotic) in the intervention groups was **0.27 standard deviations lower** (0.51 to 0.04 lower)		408 (2 studies)	⊕⊕⊖⊖ **low**[1,2]	SMD -0.27 (-0.51 to -0.04)
Depression symptomatology (switch to SSRI + antipsychotic versus switch to antipsychotic-only) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 8-12 weeks		The mean depression symptomatology (switch to SSRI + antipsychotic versus switch to antipsychotic-only) in the intervention groups was **0.44 standard deviations lower** (0.91 lower to 0.03 higher)		595 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,9,10]	SMD -0.44 (-0.91 to 0.03)
Depression symptomatology (switch to SSRI + antipsychotic versus switch to SSRI-only) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 8-12 weeks		The mean depression symptomatology (switch to SSRI + antipsychotic versus switch to SSRI-only) in the intervention groups was **0.13 standard deviations lower** (0.35 lower to 0.1 higher)		591 (2 studies)	⊕⊕⊖⊖ **low**[1,2]	SMD -0.13 (-0.35 to 0.1)
Discontinuation for any reason (switch to SSRI versus switch to non-SSRI	**Study population**		**RR 0.86** (0.65 to 1.16)	718 (3 studies)	⊕⊖⊖⊖ **very low**[2,11,12]	
	217 per 1000	**187 per 1000** (141 to 252)				

Update 2017

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	Switch to another antidepressant/non-antidepressant agent (head-to-head)				
AD) Number of participants discontinuing for any reason (including adverse events) Follow-up: 4-12 weeks	**Moderate**					
	202 per 1000	**174 per 1000** (131 to 234)				
Discontinuation for any reason (switch to SSRI versus switch to antipsychotic) Number of participants discontinuing for any reason (including adverse events) Follow-up: 8-12 weeks	**Study population**		**RR 0.82** (0.56 to 1.18)	408 (2 studies)	⊕⊖⊖⊖ **very low**[2,11,12]	
	243 per 1000	**199 per 1000** (136 to 286)				
	Moderate					
	256 per 1000	**210 per 1000** (143 to 302)				
Discontinuation for any reason (switch to SNRI versus switch to atypical antidepressant) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 8 weeks	**Study population**		**RR 0.99** (0.44 to 2.24)	105 (1 study)	⊕⊖⊖⊖ **very low**[2,13,14]	
	182 per 1000	**180 per 1000** (80 to 407)				
	Moderate					
	182 per 1000	**180 per 1000** (80 to 408)				
Discontinuation for any reason (switch to SSRI + antipsychotic versus switch to antipsychotic-only) Number of participants discontinuing for any reason (including adverse events) Follow-up: 8-12 weeks	**Study population**		**RR 0.89** (0.65 to 1.21)	595 (2 studies)	⊕⊖⊖⊖ **very low**[2,11,12]	
	243 per 1000	**216 per 1000** (158 to 294)				
	Moderate					
	256 per 1000	**228 per 1000** (166 to 310)				
	Study population					

Update 2017

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	Switch to another antidepressant/non-antidepressant agent (head-to-head)				
Discontinuation for any reason (switch to SSRI + antipsychotic versus switch to SSRI-only) Number of participants discontinuing for any reason (including adverse events) Follow-up: 8-12 weeks	**198 per 1000**	**222 per 1000** (154 to 315)	**RR 1.12** (0.78 to 1.59)	591 (2 studies)	⊕⊖⊖⊖ **very low**[2,11,15]	
	Moderate					
	199 per 1000	**223 per 1000** (155 to 316)				
Discontinuation due to adverse events (switch to SSRI versus switch to non-SSRI AD) Number of participants discontinuing due to adverse events Follow-up: 4-14 weeks	**Study population**		**RR 0.87** (0.66 to 1.14)	1253 (3 studies)	⊕⊖⊖⊖ **very low**[2,11,14]	
	179 per 1000	**156 per 1000** (118 to 204)				
	Moderate					
	82 per 1000	**71 per 1000** (54 to 93)				
Discontinuation due to adverse events (switch to SSRI versus switch to antipsychotic) Number of participants discontinuing due to adverse events Follow-up: 8-12 weeks	**Study population**		**RR 0.39** (0.16 to 0.91)	408 (2 studies)	⊕⊖⊖⊖ **very low**[2,7,11]	
	92 per 1000	**36 per 1000** (15 to 84)				
	Moderate					
	89 per 1000	**35 per 1000** (14 to 81)				
Discontinuation due to adverse events (switch to SNRI versus switch to atypical antidepressant) Number of participants discontinuing due to adverse events Follow-up: 8-14 weeks	**Study population**		**RR 0.78** (0.57 to 1.07)	589 (2 studies)	⊕⊖⊖⊖ **very low**[2,11,14]	
	225 per 1000	**175 per 1000** (128 to 241)				
	Moderate					
	136 per 1000	**106 per 1000** (78 to 146)				

Update 2017

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	Switch to another antidepressant/non-antidepressant agent (head-to-head)				
Discontinuation due to adverse events (switch to SSRI + antipsychotic versus switch to antipsychotic-only) Number of participants discontinuing due to adverse events Follow-up: 8-12 weeks	**Study population**		**RR 0.98** (0.48 to 2.03)	595 (2 studies)	⊕⊖⊖⊖ **very low**[2,11,14]	
	92 per 1000	**90 per 1000** (44 to 187)				
	Moderate					
	89 per 1000	**87 per 1000** (43 to 181)				
Discontinuation due to adverse events (switch to SSRI + antipsychotic versus switch to SSRI-only) Number of participants discontinuing due to adverse events Follow-up: 8-12 weeks	**Study population**		**RR 2.41** (1.07 to 5.42)	591 (2 studies)	⊕⊖⊖⊖ **very low**[2,7,11]	
	35 per 1000	**84 per 1000** (37 to 188)				
	Moderate					
	39 per 1000	**94 per 1000** (42 to 211)				

Update 2017

Notes:
[1] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration and outcome assessment
[2] Data cannot be extracted or is not reported for all outcomes and/or funding from pharmaceutical companies
[3] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[4] Unclear (or high risk) randomisation method and unclear method of allocation concealment, and unclear blinding of intervention administrator(s)
[5] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[6] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[7] Events<300
[8] I-squared>50%
[9] I-squared>80%
[10] 95% CI crosses both line of no effect and threshold for clinically important benefit (SMD -0.5)
[11] Unclear randomisation method and method of allocation concealment and unclear blinding of intervention administrator(s)
[12] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 0.75)
[13] High risk of bias associated with randomisation method due to significant difference between groups at baseline and unclear blinding of intervention administrator(s)
[14] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[15] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)

Table 142: Study information table for trials included in the meta-analysis of switching to a combined psychological and pharmacological intervention versus switching to a psychological intervention-only

	CBT individual (under 15 sessions) + antipsychotic versus CBT individual (under 15 sessions)-only
Total no. of studies (N randomised)	1 (22)
Study ID	Chaput 2008
Country	Canada
Diagnostic status	DSM-IV unipolar major depression
Age range (mean)	Range NR (43.3)
Sex (% female)	73
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR
Mean months (SD) since onset of current episode	22.5 (14.7)
No. (SD) of previous depressive episodes	NR
Details of inadequate response/treatment resistance	TRD: Failure of 2 (or more) 8-week treatments with 2 different classes of antidepressants and for at least 3 of those eight weeks, doses were required to be at or near the highest therapeutically recommended doses (verified by examining any pertinent medical records or charts) plus failure to respond (< 40% reduction or a score >18 on the HAMD) to lithium augmentation (open-label lithium augmentation [≥ 600 mg per day, serum levels of between 0.6 and 0.9 mEq/L by day 7]) of AD treatment in a 3-week prospective treatment phase
Augmented/previous treatment	Previous treatment: Lithium augmentation of AD
Baseline severity	MADRS 30.2 (More severe)
Intervention details (mean dose)	CBT individual 12x weekly 1-hour sessions (mean attended 11 sessions [SD=2]) + quetiapine 25-400mg/day (mean final dose 147.7mg [SD=112 mg])
Comparator details (mean dose, if applicable)	CBT individual 12x weekly 1-hour sessions (mean attended 7 sessions [SD=5]) + placebo 25-400mg/day (mean final dose 209.1mg [SD=120 mg])
Treatment length (weeks)	12
Notes: Abbreviations: mg=milligram, NR=not reported, SD=standard deviation, TRD=treatment-resistant depression	

Update 2017

Table 143: Summary of findings table for switching to a combined psychological and pharmacological intervention versus switching to a psychological intervention-only

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	Switch to combined psych and pharm intervention versus switch to psych intervention-only				
Discontinuation for any reason (CBT individual [under 15 sessions] + antipsychotic versus CBT individual [under 15 sessions]-only) Number of participants discontinuing for any reason (including adverse events) Follow-up: mean 12 weeks	**Study population**		**RR 0.17** (0.02 to 1.17)	22 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	545 per 1000	**93 per 1000** (11 to 638)				
	Moderate					
	546 per 1000	**93 per 1000** (11 to 639)				

Notes:
[1] Unclear randomisation method and unclear blinding of intervention administrator(s)
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 0.75)
[3] Efficacy data cannot be extracted and study funded by pharmaceutical company

Update 2017

8.4 Economic evidence

The systematic search of the literature identified 8 studies on the cost effectiveness of interventions for the management of adults with depression that failed to respond to previous treatment. Of these, 2 were UK studies assessing psychological interventions (Scott et al., 2003; Hollinghurst et al., 2014 and Wiles et al., 2016) and 2 were UK studies assessing pharmacological interventions (Benedict et al., 2010; Edwards et al.,2013). Following the hierarchy of inclusion criteria regarding country settings, one Swedish study (Nordström et al., 2010) and 3 US studies (Olgiati et al., 2013; Malone, 2007; Taneja et al., 2012) that assessed the cost effectiveness of pharmacological interventions in adults with depression that failed to respond to previous treatment were also included in the review, since they assessed interventions that had not been evaluated in UK studies. Details on the methods used for the systematic search of the economic literature, including inclusion criteria for each review question, are described in Chapter 3. Full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix Q. Completed methodology checklists of the studies are provided in Appendix P. 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 R.

8.4.1 Psychological interventions

Scott and colleagues (2003) conducted a cost effectiveness analysis alongside a RCT (Paykel1999; N=158) that compared cognitive therapy in addition to antidepressant therapy

and clinical management versus antidepressant therapy and clinical management alone, in adults who were in an episode of major depression within the past 18 months but not in the past 2 months, and who had residual symptoms over at least 8 weeks (HAMD ≥ 8 and BDI ≥ 9). The perspective of the analysis was that of the NHS and personal social services (PSS). Healthcare cost elements consisted of interventions (cognitive therapy, medication, clinical management), inpatient care, day hospital, staff time (GP, social worker, community psychiatric nurse, therapist/counsellor), group therapy and marital therapy. National and local inpatient unit costs were used. The outcome measure was the percentage of relapses prevented. The duration of the analysis was 17 months.

Cognitive therapy in addition to antidepressants and clinical management was significantly more effective and more costly than antidepressant therapy and clinical management alone, with an Incremental Cost Effectiveness Ratio (ICER) of £7,030/additional relapse prevented (2015 prices). This figure was higher depending on the method of imputation of missing data and reached £11,462 when a complete case analysis, using 65% of participants, was conducted. The probability of cognitive therapy in addition to antidepressant being cost-effective was 0.60 and 0.80 at a willingness to pay (WTP) of £9,700 and £13,800 per relapse prevented, respectively. This probability was sensitive to the method of missing data imputation. The study is partially applicable to the NICE decision-making context as it does not use the QALY as the measure of outcome and interpretation of the results requires judgement as to whether the additional unit of benefit (prevention of one relapse) is worth the additional cost of £7,030. The study is characterised by minor limitations.

Hollinghurst and colleagues (2014) conducted a cost consequence and cost-utility analysis alongside a RCT (Wiles2013; N=469) to assess the cost effectiveness of CBT in addition to TAU versus TAU alone, in adults with major depression who had adhered to antidepressant medication for at least 6 weeks in primary care, but who continued to have significant depressive symptoms (BDI-II score ≥14 and ICD-10 diagnosis of depression), in the UK; TAU comprised GP care, including antidepressant treatment as judged appropriate by the person's GP or a referral, as required. The time horizon of the analysis was 12 months; 3-5 year follow up data were reported in a separate publication (Wiles et al., 2016). The perspective of the cost-utility analysis was that of the NHS and PSS, with cost elements comprising intervention (CBT), medication, primary and community mental and general health care, and specialist (secondary) mental health care. National unit costs were used. A number of outcomes were assessed, such as the change in BDI-II score, response and remission rates, and the SF-12 mental and physical subscales. QALYs were estimated using the EQ-5D (UK tariff), with SF-6D ratings being used for the estimation of QALYs in a sensitivity analysis.

CBT was found to be associated with a significant increase in total NHS and PSS costs and was also significantly better than control in a number of outcomes including response, the SF-12 mental sub-scale score and the QALY, both at 12 months and at the 3-5 year follow up. At 12 months, the ICER of CBT plus TAU versus TAU alone was £16,271/QALY (2015 prices). The probability of CBT being cost-effective was 0.74 and 0.91 at the NICE lower and upper cost effectiveness threshold of £20,000 and £30,000/QALY, respectively. Results were not sensitive to a change in psychologist unit costs and to the exclusion of hospitalisation costs; in contrast, results were sensitive to estimation of QALYs using the SF-6D instead of EQ-5D, with the ICER rising at £32,328/QALY. Analysis of participants with full complete data (instead of imputation of missing data) resulted in ICER of £20,036/QALY. At the 3-5 year follow up, the ICER of CBT versus TAU dropped at £5,482/QALY (2015 prices) with the probability of CBT being cost-effective rising at 0.92 and 0.95, at the NICE lower and upper cost effectiveness threshold of £20,000 and £30,000/QALY, respectively. The study is directly applicable to the NICE decision-making context and is characterised by minor limitations.

Update 2017

8.4.21 Pharmacological interventions

Benedict and colleagues (2010) constructed an economic model to evaluate the cost effectiveness of duloxetine, venlafaxine and mirtazapine in adults with severe major depression who failed previous SSRI treatment and were referred to mental health specialists in secondary care in the UK. The duration of the analysis was 48 weeks. The analysis adopted the perspective of the Scottish NHS, with costs including medication, A&E visits, staff time (GPs, psychiatrists) and hospitalisation. Resource use estimates were based on expert opinion; national unit costs were used. The outcome measure was the QALY, based on EQ-5D ratings (UK tariff). Efficacy data were obtained from meta-analyses of RCTs, with randomisation rules possibly being broken. Duloxetine was found to dominate both venlafaxine and mirtazapine and to have a probability of being cost-effective of 0.80 at the NICE lower cost effectiveness threshold of £20,000/QALY. Although the study is directly applicable to the NICE decision-making context, it is characterised by potentially serious limitations, including the methods for meta-analysis and evidence synthesis (selective use of RCTs and synthesis that appears to have potentially broken randomisation) and the fact that it was funded by industry, which may have introduced bias in the analysis.

Edwards and colleagues (2013) developed an economic model to assess the cost-utility of atypical antipsychotics versus lithium, both as adjuncts to an SSRI, for the treatment of adults with treatment-resistant depression in the UK. The study adopted a NHS and PSS perspective and considered medication costs, healthcare professional time (GP, community mental health teams, crisis resolution and home treatment teams), hospitalisation and monitoring (laboratory testing) costs. Efficacy data were taken from a systematic review and network meta-analysis that enabled an indirect comparison between the two interventions, using 6 RCTs comparing olanzapine plus fluoxetine versus fluoxetine alone in people with treatment-resistant depression and 1 RCT comparing lithium plus fluoxetine versus fluoxetine alone in people who had failed at least one antidepressant; a common class effect was assumed for SSRIs and also for antipsychotics. It needs to be noted that data on lithium as adjunct to an SSRI were taken from a population that had failed to respond to one previous SSRI (and not from people with treatment-resistant depression) due to lack of more relevant data. In order to estimate the effect of each intervention, a fixed baseline MADRS score was assumed for both arms; the change in MADRS scores at endpoint was assumed to have a normal distribution, which was used to estimate proportions of people in the remission, response and no response states.

Resource use estimates were mainly based on clinical expert opinion, with the exception of the length of hospitalisation, which was based on national hospital episode statistics. In order to estimate medication costs in each arm of the model, it was assumed, based on expert advice, that antipsychotic use comprised 30% aripiprazole, 30% olanzapine, 20% quetiapine, and 20% risperidone; and SSRI use comprised 20% citalopram, 20% escitalopram, 30% fluoxetine, and 30% sertraline. The study utilised national unit costs. The outcome measure was the QALY estimated based on EQ-5D ratings (UK tariff). The time horizon of the analysis was 12 months.

Augmentation of SSRIs with lithium was found to dominate augmentation of SSRIs with an antipsychotic; the probability of lithium being dominant versus antipsychotics (both as adjuncts to an SSRI) was 1. Results were sensitive to the efficacy of augmentation strategies and discontinuation rates; they were robust under different assumptions regarding resource use, as well as under changes in remission and relapse risk at follow-up. The study is directly applicable to the UK context and is characterised by potentially serious limitations, comprising mainly the source of efficacy data (i.e. the lack of evidence on treatment-resistant depression treated with lithium as an adjunct on a SSRI), the assumptions made around baseline and endpoint MADRS scores, and the fact that all resource use was based on expert opinion.

Update 2017

Nordström and colleagues (2010) developed an economic model to evaluate the cost effectiveness of escitalopram, duloxetine and venlafaxine in adults with major depression treated in primary care, who had had a history of treatment with another antidepressant within the previous 6 months, in Sweden. The time horizon of the analysis was 6 months. The analysis adopted a societal perspective but healthcare costs were reported separately and included medication, staff time (GP, psychiatrist, other doctors e.g. neurologist, cardiologist, psychotherapist, counsellor, psychologist, nurse), hospitalisation and treatment of side effects. Resource use estimates were based on a cohort study conducted in 56 primary care centres in Sweden over 6 months; national unit costs were used. The outcome measure was the probability of remission (defined as a MADRS total score ≤ 12) achieved after 8 weeks of treatment and sustained until the end of 6 months; and the QALY estimated based on EQ-5D ratings (UK tariff). Efficacy data were derived from pooled analysis of trial data, including only participants who had already received antidepressant therapy prior to randomisation; data for duloxetine and venlafaxine were pooled together. Considering only healthcare costs, escitalopram was found to dominate both duloxetine and venlafaxine and to have a probability of being cost-effective of more than 0.98 at the NICE lower cost effectiveness threshold of £20,000/QALY. The study is only partially applicable to the NICE decision-making context and is characterised by potentially serious limitations, including the methods for evidence synthesis (selective use of RCTs and data pooling for two of the assessed interventions) and the fact that it was funded by industry, which may have introduced bias in the analysis.

The other 3 studies included in the economic literature review assessed different pharmacological treatment options in adults with depression who responded inadequately to previous treatment using decision-analytic economic modelling. All 3 studies were conducted in the US. Olgiati and colleagues (2013) compared different strategies for adults with depression that did not remit following pharmacological treatment (citalopram), comprising continuation of current treatment (citalopram), switching to sertraline or venlafaxine, or augmentation of citalopram with bupropion. The study reported that both switching and augmentation strategies were more cost-effective than continuation of current treatment. However, efficacy data for the 3 strategies were taken from different studies without using a common comparator, thus breaking randomisation rules. Malone (2007) compared different SSRIs (including generic SSRIs, escitalopram, paroxetine controlled release, sertraline and venlafaxine) in adults with major depression who failed to achieve remission with previous treatment with SSRIs. The study reported that paroxetine controlled release and sertraline were dominated by other antidepressant options. Efficacy estimates were based on a review of published trial data and further assumptions; evidence synthesis was done by naïve addition of efficacy data, leading to breaking of randomisation rules; the study was funded by industry, which may have introduced further bias to the analysis. Finally, Taneja and colleagues (2012) compared different antipsychotics (aripiprazole, quetiapine and olanzapine) as adjuncts to antidepressants versus antidepressant treatment alone, in adults with major depression who had responded inadequately to previous antidepressant therapy. Efficacy data were derived from a meta-analysis of published phase III clinical trials and indirect comparison using placebo as baseline comparator. The study found that quetiapine as an adjunct to antidepressants and the combination of olanzapine/fluoxetine were extendedly dominated and the ICER of aripiprazole as an adjunct to antidepressants versus antidepressants alone was £2,555 per person responding (converted and uplifted to 2015 UK pounds). The time horizon was too short (only 6 weeks) to allow assessment of the cost effectiveness of interventions over the duration of the depressive episode; moreover, the study was funded by industry, which may have introduced additional bias in the analysis. All 3 US studies are partially applicable to the UK context and all are characterised by very serious limitations. Therefore, they have not been considered further when formulating recommendations.

Update 2017

8.5 Clinical evidence statements

8.5.1 Dose escalation strategies

- Very low quality evidence from 2-3 studies (N=270-327), suggests that there is no clinically important or statistically significant benefit of increasing the dose of an SSRI, relative to continuing at the same dose of the SSRI, on the rate of remission or the rate of response (as measured by the number of participants rated as much or very much improved on the CGI-I) in adults with depression who have responded inadequately to previous treatment. Evidence from 1 of these studies (N=57) suggests the same pattern of results for depression symptomatology. In fact, very low quality evidence from 2 studies (N=252) suggests that there is a statistically significant benefit in favour of continuing at the same dose of an SSRI, relative to increasing the dose, on a different measure of response (the number of participants showing at least 50% improvement from baseline on the HAM-D). Very low quality evidence from 1-5 studies (N=60-332) suggests that there are no clinically important or statistically significant harms associated with increasing the dose of an SSRI as measured by discontinuation for any reason and discontinuation due to adverse events, conversely, there was some suggestion of higher drop-out in the same dose arm (although absolute numbers are small).
- Very low quality single-study evidence (N=472) suggests a small, but statistically significant and potentially clinically important, benefit of increasing the dose of a continued SSRI (escitalopram), relative to switching to an SNRI (duloxetine), on the rate of remission in adults with depression who have responded inadequately to previous treatment. However, the same study found neither clinically important nor statistically significant effects on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS or the number of participants rated as much or very much improved on the CGI-I), or on depression symptomatology. This study found no evidence for clinically important or statistically significant harms associated with increasing the dose of an SSRI as measured by discontinuation for any reason or due to adverse events.
- Very low quality evidence from 2 studies (N=94) suggests a clinically important, but not statistically significant, benefit of increasing the dose of an SSRI (fluoxetine), relative to TCA (desipramine) augmentation of fluoxetine (at the lower continued dose), on the rate of remission and on depression symptomatology in adults with depression who have responded inadequately to previous treatment. Evidence from 1-2 of these studies (N=27-94) suggests no clinically important or statistically significant harms associated with increasing the dose of an SSRI as measured by discontinuation for any reason or discontinuation due to adverse events, conversely, there was some suggestion of higher drop-out in the same dose arm (although absolute numbers are small).
- Very low quality evidence from 2 studies (N=96) suggests a clinically important and statistically significant benefit of increasing the dose of an SSRI (fluoxetine), relative to lithium augmentation of fluoxetine (at the lower continued dose), on the rate of remission in adults with depression who have responded inadequately to previous treatment. The same two studies found a trend for the same pattern of results on depression symptomatology. There was no evidence from these 2 studies for clinically important or statistically significant harms associated with increasing the dose of an SSRI as measured by discontinuation for any reason or discontinuation due to adverse events, conversely, there was some suggestion of higher drop-out in the same dose arm (although absolute numbers are small).
- Very low quality single-study evidence (N=195) suggests a clinically important and statistically significant benefit in favour of TeCA (mianserin) augmentation of an SSRI (sertraline) at the lower continued dose, relative to increasing the dose of sertraline, on the rate of remission in adults with depression who have responded inadequately to previous treatment. The same study found neither clinically important nor statistically

Update 2017

significant effects on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or the number of participants rated as much or very much improved on the CGI-I) or on discontinuation for any reason.

8.5.25 Augmentation strategies

- Low quality evidence from 20 studies (N=3110) suggests a clinically important and statistically significant benefit of augmenting the antidepressant with any active agent (atypical antidepressant, antipsychotic, lithium, lamotrigine, omega-3 fatty acid or methylphenidate) relative to augmentation with placebo, on the rate of response as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or MADRS, in adults with depression who have responded inadequately to previous treatment. Very low quality evidence from 5 studies (N=257) suggests a clinically important but not statistically significant benefit of augmentation with any active agent (mirtazapine, lithium, lamotrigine, buspirone or methylphenidate) relative to placebo on response as measured by the number of participants rated as much or very much improved on the CGI-I.
- Very low quality evidence from 3 studies (N=531) suggests a small but statistically significant benefit of augmenting an SSRI with any active agent (mianserin, olanzapine or ziprasidone), relative to continuing with the SSRI-only, on depression symptomatology.
- Very low quality evidence from 2 studies (N=86) suggests a clinically important and statistically significant benefit of augmenting the antidepressant with an atypical antidepressant (mirtazapine or bupropion), relative to augmentation with placebo, on the rate of remission in adults with depression who have responded inadequately to previous treatment. Evidence from 1 of these studies (N=26) suggests the same pattern of results with mirtazapine as an augmentation agent on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D and by the number of participants rated as much or very much improved on the CGI-I) and on depression symptomatology. Evidence from these same studies suggests neither clinically important nor statistically significant harms associated with atypical antidepressant augmentation as measured by discontinuation for any reason and discontinuation due to adverse events, conversely, there is some suggestion of higher drop-out in the placebo arm (although absolute numbers are small).
- Low quality evidence from 9-10 studies (N=2581-2604) suggests a clinically important and statistically significant benefit of augmenting the antidepressant with an antipsychotic (aripiprazole, quetiapine or risperidone), relative to augmentation with placebo, on the rate of remission and response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or MADRS) in adults with depression who have responded inadequately to previous treatment. Very low quality evidence from 3 of these studies (N=462) also suggests a trend for the same pattern of results on depression symptomatology. Low to very low quality evidence from all 10 studies (N=2706) suggests a statistically significant, or clinically important and statistically significant, harm associated with antipsychotic augmentation as measured by discontinuation for any reason and discontinuation due to adverse events respectively.
- Low quality evidence from 2 studies (N=461) suggests a small but statistically significant benefit of antipsychotic (olanzapine or ziprasidone) augmentation of an SSRI, relative to continuing with the SSRI-only, on depression symptomatology in adults with depression who have responded inadequately to previous treatment. However, very low quality evidence from 3 studies (N=551) suggests neither clinically important nor statistically significant benefits of antipsychotic (olanzapine, risperidone or ziprasidone) augmentation of an SSRI, relative to continuing with the SSRI-only, on the rate of remission or the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or MADRS). There is also evidence from 2 of these studies (N=461-467) for clinically important and statistically significant harm as

Update 2017

measured by discontinuation due to any reason and discontinuation due to adverse events.

- Very low quality evidence from 3-4 studies (N=76-110) suggests a clinically important and statistically significant benefit of augmenting the antidepressant with lithium, relative to augmentation with placebo, on the rate of remission in adults with depression who have responded inadequately to previous treatment, and a clinically important but not statistically significant benefit on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D). However, very low quality evidence from 1-3 (N=35-83) of these studies suggests neither clinically important nor statistically significant benefits on depression symptomatology or on a different measure of response (the number of participants rated as much or very much improved on the CGI-I). Very low quality evidence from 5 studies (N=165) suggests a clinically important, but not statistically significant, harm associated with lithium augmentation as measured by discontinuation due to adverse events, however, absolute numbers are small. Effects on discontinuation for any reason (K=6; N=200) were neither clinically important nor statistically significant.
- Very low quality single-study evidence (N=24) suggests a clinically important and statistically significant benefit of augmenting an SSRI (citalopram) with lithium, relative to continuing with citalopram-only, on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) in adults with depression who have responded inadequately to previous treatment.
- Low quality single-study evidence (N=33) suggests a clinically important and statistically significant benefit of augmenting the antidepressant with a thyroid hormone (T3) relative to placebo augmentation on depression symptomatology in adults with depression who have responded inadequately to previous treatment, and a clinically important but not statistically significant benefit of T3 relative to placebo augmentation on the rate of remission. Low quality evidence from 2 studies (N=51) suggests neither clinically important nor statistically significant harms associated with T3 augmentation as measured by discontinuation for any reason or due to adverse events with no drop-out in either arm, although relative risk is not estimable and sample size is small.
- Very low quality single-study evidence (N=93) suggests a clinically important but not statistically significant benefit of augmentation of an SSRI (paroxetine) with a thyroid hormone, relative to continuing with paroxetine-only, on the rate of remission and the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) in adults with depression who have responded inadequately to previous treatment.
- Very low quality single-study evidence (N=60) suggests a clinically important, but not statistically significant, benefit of augmenting the antidepressant with a stimulant (methylphenidate) relative to placebo augmentation on the rate of remission and the rate of response (as measured by the number of participants rated as much or very much improved on the CGI-I) in adults with depression who have responded inadequately to previous treatment. However, very low quality evidence from 2 studies (N=205) suggests neither a clinically important nor statistically significant benefit of augmenting the antidepressant with methylphenidate relative to placebo on a different measure of response (the number of participants showing at least 50% improvement from baseline on the HAM-D or MADRS) or on depression symptomatology. Evidence from these same studies also suggests a clinically important, but not statistically significant, harm associated with methylphenidate augmentation as measured by discontinuation due to any reason and discontinuation due to adverse events.
- Very low quality evidence from 2 studies (N=130) suggests neither a clinically important nor statistically significant benefit of augmenting the antidepressant with an anticonvulsant (lamotrigine), relative to augmentation with placebo, on the rate of response as measured by the number of participants showing at least 50% improvement from baseline on the MADRS, or on depression symptomatology, in adults with depression who have

Update 2017

responded inadequately to previous treatment. Evidence from 1 of these studies (N=34) suggests a clinically important, but not statistically significant, benefit in favour of placebo augmentation of the antidepressant, relative to lamotrigine augmentation, on a different measure of response (the number of participants rated as much or very much improved on the CGI-I). Evidence from both of these studies suggests neither clinically important nor statistically significant harms associated with lamotrigine augmentation as measured by discontinuation due to any reason and discontinuation due to adverse events.

- Very low quality single-study evidence (N=84) suggests neither a clinically important nor statistically significant benefit, of augmenting an SSRI (paroxetine) with an anticonvulsant (sodium valproate) relative to continuing with paroxetine-only, on the rate of remission or the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) in adults with depression who have responded inadequately to previous treatment.
- Very low quality single-study evidence (N=69) suggests neither a clinically important nor statistically significant benefit of augmenting the antidepressant with an omega-3 fatty acid, relative to augmentation with placebo, on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS) in adults with depression who have responded inadequately to previous treatment. Evidence from this same study suggests neither clinically important nor statistically significant harms associated with omega-3 augmentation as measured by discontinuation for any reason or discontinuation due to adverse events, conversely, there was some suggestion of higher drop-out in the placebo arm (although absolute numbers are small).
- Very low quality single-study evidence (N=102) suggests neither a clinically important nor statistically significant benefit of augmenting the antidepressant with an anxiolytic (buspirone), relative to augmentation with placebo, on the rate of response (as measured by the number of participants rated as much or very much improved on the CGI-I) in adults with depression who have responded inadequately to previous treatment. Evidence from this same study suggests neither clinically important nor statistically significant harms associated with buspirone augmentation as measured by discontinuation for any reason or discontinuation due to adverse events, conversely, there was some suggestion of higher drop-out (for any reason) in the placebo arm.
- Very low quality single-study evidence (N=91) suggests a clinically important, but not statistically significant, benefit in favour of continuing with paroxetine-only relative to buspirone augmentation of paroxetine on the rate of remission in adults with depression who have responded inadequately to previous treatment. Evidence from the same study suggests neither a clinically important nor a statistically significant benefit of buspirone augmentation of paroxetine, relative to continuing with paroxetine-only, on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D).
- Very low quality single-study evidence (N=70) suggests a moderate and statistically significant benefit of augmenting an SSRI with a TeCA (mianserin) relative to continuing with the SSRI-only, on depression symptomatology in adults with depression who have responded inadequately to previous treatment, and very low quality evidence from 2 studies (N=266) suggests a clinically important, but not statistically significant, benefit of mianserin augmentation on the rate of remission. However, evidence from the same two studies suggests neither a clinically important nor statistically significant benefit on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or the number of participants rated as much or very much improved on the CGI-I). There is also evidence from these same studies for a clinically important, but not statistically significant, harm as measured by discontinuation for any reason and discontinuation due to adverse events.
- Very low quality single-study evidence (N=92) suggests neither a clinically important nor statistically significant benefit of augmentation of an SSRI (paroxetine) with a SARI (trazodone), relative to continuing with paroxetine-only, on the rate of remission or the rate

Update 2017

of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) in adults with depression who have responded inadequately to previous treatment.

- In head-to-head comparisons of lithium compared with other pharmacological augmentation agents, there was very low quality evidence from 4-7 studies (N=774) for a statistically significant benefit in favour of 'other' augmentation agents (desipramine, quetiapine, T3 or lamotrigine) relative to lithium on the rate of remission in adults with depression who have responded inadequately to previous treatment. While, very low quality evidence from 4-5 studies (N=304-646) suggested neither clinically important nor statistically significant differences between lithium and 'other' augmentation agents (quetiapine/desipramine, T3 or lamotrigine) on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D, MADRS or QIDS) or depression symptomatology. Very low quality evidence from 7 studies (N=662-736) suggested a clinically important, but not statistically significant, harm associated with lithium relative to other augmentation agents as measured by discontinuation for any reason and discontinuation due to adverse events.
- Very low quality evidence from 2 studies (N=94) suggests no significant difference between lithium and desipramine as augmentation agents (of fluoxetine) on the rate of remission, depression symptomatology or discontinuation for any reason, in adults with depression who have responded inadequately to previous treatment. Evidence from 1 of these studies (N=26) suggests a clinically important, but not statistically significant, harm associated with desipramine relative to lithium augmentation, as measured by discontinuation due to adverse events. However, this is a small single study and absolute numbers are small.
- Very low quality evidence from 2 studies (N=470) suggests a clinically important, but not statistically significant benefit of quetiapine relative to lithium as augmentation agents of antidepressants, on the rate of remission and discontinuation for any reason in adults with depression who have responded inadequately to previous treatment. Evidence from 1-2 of these studies (N=450-470) suggested neither clinically important nor statistically significant differences between lithium and quetiapine as augmentation agents for rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or the number of participants rated as much or very much improved on the CGI-I) or on discontinuation due to adverse events.
- Very low quality evidence from 1-2 studies (N=142-176) suggests a clinically important, but not statistically significant benefit of thyroid hormone (T3) relative to lithium as augmentation agents of antidepressants, on the rate of remission, the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the QIDS) and discontinuation for any reason, in adults with depression who have responded inadequately to previous treatment. However, evidence from these same two studies suggests neither clinically important nor statistically significant differences between lithium and T3 on depression symptomatology. Very low quality evidence from 3 studies (N=196) suggests a clinically important and statistically significant harm associated with lithium relative to T3 augmentation, as measured by discontinuation due to adverse events.
- Very low quality single-study evidence (N=93) suggests a clinically important, but not statistically significant, benefit of thyroid hormone relative to antipsychotic (risperidone) augmentation (of paroxetine) on the rate of remission in adults with depression who have responded inadequately to previous treatment. This study also found a trend for the same pattern of results on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D).
- Very low quality single-study evidence (N=94-95) suggests neither clinically important nor statistically significant benefits of thyroid hormone augmentation (of paroxetine) relative to either anxiolytic (buspirone) or SARI (trazodone) augmentation (of paroxetine) on the rate of remission or the rate of response (as measured by the number of participants showing

Update 2017

at least 50% improvement from baseline on the HAM-D) in adults with depression who have responded inadequately to previous treatment.

- Low to very low quality single-study evidence (N=34) suggests a clinically important and statistically significant benefit of an anticonvulsant (lamotrigine) relative to lithium as augmentation agents of antidepressants on depression symptomatology in adults with depression who have responded inadequately to previous treatment, a clinically important but not statistically significant benefit on the rate of remission, and a trend for the same pattern of results on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D). Evidence from this study suggested no significant difference between lithium and lamotrigine in discontinuation for any reason or discontinuation due to adverse events.
- Very low quality single-study evidence (N=84) suggests a clinically important and statistically significant benefit of an anticonvulsant (sodium valproate) relative to an antipsychotic (risperidone) as augmentation agents (of paroxetine) on the rate of remission in adults with depression who have responded inadequately to previous treatment. This study also found a trend for the same pattern of results on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D).
- Very low quality single-study evidence (N=85-87) suggests a clinically important, but not statistically significant, benefit of an anticonvulsant (sodium valproate) relative to an anxiolytic (buspirone) or a thyroid hormone as augmentation agents (of paroxetine) on the rate of remission in adults with depression who have responded inadequately to previous treatment. However, the same study found no differences between sodium valproate and either buspirone or a thyroid hormone (as augmentation agents of paroxetine) on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D).
- Very low quality single-study evidence (N=91) suggests neither a clinically important nor statistically significant difference between antipsychotic augmentation (risperidone + paroxetine) and anxiolytic augmentation (buspirone + paroxetine) on the rate of remission or the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) in adults with depression who have responded inadequately to previous treatment.
- Very low quality single-study evidence (N=92) suggests a clinically important, but not statistically significant, benefit of a SARI (trazodone) relative to an antipsychotic (risperidone) as augmentation agents (of paroxetine) on the rate of remission in adults with depression who have responded inadequately to previous treatment. This study also found a trend for the same pattern of results on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D).
- Very low quality single-study evidence (N=86-93) suggests neither a clinically important nor statistically significant benefit of SARI augmentation (trazodone + paroxetine), compared with either anticonvulsant augmentation (sodium valproate + paroxetine) or anxiolytic augmentation (buspirone + paroxetine), on the rate of remission or the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) in adults with depression who have responded inadequately to previous treatment.
- Very low quality single-study evidence (N=565) suggests a statistically significant but very small benefit of an atypical antidepressant (bupropion) relative to an anxiolytic (buspirone) augmentation (of citalopram) on depression symptomatology in adults with depression who have responded inadequately to previous treatment. However, this study found no significant difference between bupropion and buspirone on the rate of remission or the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the QIDS). Evidence from this study suggests a clinically important and statistically significant harm associated with buspirone (+ citalopram)

Update 2017

relative to bupropion (+ citalopram) as measured by discontinuation due to adverse events.

- Very low quality single-study evidence (N=173) suggests a clinically important and statistically significant benefit of MBCT relative to attention-placebo augmentation of antidepressants on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D), and a clinically important (but not statistically significant) benefit on the rate of remission, in adults with depression who have responded inadequately to previous treatment. Low quality evidence from another single study (N=43) also suggests a benefit for MBCT relative to attention-placebo on depression symptomatology. There was also very low quality evidence from both of these studies (N=223) suggesting lower discontinuation for any reason in the MBCT arm, although this effect is not statistically significant.
- Low to very low quality evidence from 4 studies (N=869) suggests clinically important and statistically significant benefits of augmenting antidepressant treatment with a cognitive behavioural therapy (CBASP or CBT), relative to continuing with the antidepressant only, on the rate of remission and depression symptomatology in adults with depression who have responded inadequately to previous treatment. Evidence from 2 of these studies (N=461) suggests the same pattern of results on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or BDI). Very low quality evidence from all 4 studies (N=965) suggests neither clinically important nor statistically significant harms associated with cognitive behavioural therapy augmentation as measured by discontinuation for any reason.
- Very low quality evidence from 3 studies (N=495) suggests a clinically important and statistically significant benefit of augmenting antidepressant treatment with any psychological intervention, relative to continuing with the antidepressant only, on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or BDI) in adults with depression who have responded inadequately to previous treatment.
- Very low quality single-study evidence (N=250) suggests a small but statistically significant benefit of augmenting antidepressant treatment with CBASP, relative to continuing with the antidepressant only, on depression symptomatology in adults with depression who have responded inadequately to previous treatment. However, the same study found neither a clinically important nor statistically significant benefit of CBASP augmentation on the rate of remission. Evidence from this study (N=296) suggests neither clinically important nor statistically significant harms associated with CBASP augmentation as measured by discontinuation for any reason or discontinuation due to adverse events, conversely, there was some suggestion of higher drop-out in the antidepressant-only arm although this effect is not statistically significant.
- Very low quality evidence from 2 studies (N=577) suggests a clinically important and statistically significant benefit of augmenting antidepressant treatment with individual CBT of 15 sessions or more, relative to continuing with the antidepressant only, on the rate of remission in adults with depression who have responded inadequately to previous treatment. Evidence from 1 of these studies (N=419) suggests the same pattern of results on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the BDI). Evidence from both these studies (N=577) also suggests a trend for the same pattern of results on depression symptomatology. However, low quality evidence from these 2 studies suggests a clinically important, but not statistically significant, harm associated with high-intensity CBT augmentation as measured by discontinuation for any reason.
- Low quality single-study evidence (N=42) suggests clinically important and statistically significant benefits of augmenting antidepressant treatment with individual CBT of 15 sessions or less, relative to continuing with the antidepressant only, on the rate of remission, the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) and depression symptomatology, in

Update 2017

adults with depression who have responded inadequately to previous treatment. Very low quality evidence from this same study suggests neither clinically important nor statistically significant harms associated with low-intensity CBT augmentation as measured by discontinuation for any reason, conversely, there was some suggestion of higher drop-out in the antidepressant-only arm although this effect is not statistically significant.

- Low to very low quality single-study evidence (N=34) suggests clinically important, but not statistically significant, benefits of augmenting antidepressant treatment with IPT relative to continuing with the antidepressant only on the rate of remission, the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) and depression symptomatology, in adults with depression who have responded inadequately to previous treatment. Very low quality evidence from this same study (N=40) suggests a clinically important, but not statistically significant, harm associated with IPT augmentation as measured by discontinuation for any reason.
- Very low quality evidence from 2 studies (N=304) suggests a clinically important, but not statistically significant, benefit of augmenting antidepressant treatment with short-term psychodynamic psychotherapy relative to continuing with the antidepressant only on the rate of remission in adults with depression who have responded inadequately to previous treatment. However, evidence from 1 of these studies (N=244) suggests neither a clinically important nor statistically significant benefit of short-term psychodynamic psychotherapy augmentation on depression symptomatology. Evidence from 1-2 of these studies (N=291-351) suggests neither a clinically important nor statistically significant harm associated with short-term psychodynamic psychotherapy augmentation as measured by discontinuation due to any reason or discontinuation due to adverse events, conversely, there was some suggestion of higher drop-out due to adverse events in the antidepressant-only arm although this effect is not statistically significant.
- Very low quality single-study evidence (N=129) suggests a clinically important, but not statistically significant, benefit of augmenting antidepressant treatment with long-term psychodynamic psychotherapy relative to continuing with the antidepressant only on the rate of remission in adults with depression who have responded inadequately to previous treatment. However, evidence from the same study suggests neither a clinically important nor statistically significant benefit of long-term psychodynamic psychotherapy augmentation on depression symptomatology. Evidence from this study suggests neither a clinically important nor statistically significant harm associated with long-term psychodynamic psychotherapy augmentation as measured by discontinuation due to any reason.
- Low quality single-study evidence (N=90) suggests neither a clinically important nor statistically significant benefit of augmenting antidepressant treatment with cognitive bibliotherapy, relative to continuing with the antidepressant only, on depression symptomatology in adults with depression who have responded inadequately to previous treatment. Very low quality evidence from this same study suggests a clinically important, but not statistically significant, harm associated with cognitive bibliotherapy augmentation as measured by discontinuation for any reason.
- Very low quality single-study evidence (N=344) suggests neither a clinically important nor statistically significant benefit of augmenting antidepressant treatment with mutual peer support, relative to continuing with the antidepressant only, on depression symptomatology in adults with depression who have responded inadequately to previous treatment. Evidence from the same study (N=387) suggests neither a clinically important nor statistically significant harm associated with mutual peer support augmentation as measured by discontinuation due to any reason.
- Low to very low quality single study evidence (N=44) suggests a clinically important and statistically significant benefit of augmenting antidepressant treatment with lithium relative to augmenting antidepressant treatment with individual CBT of less than 15 sessions on depression symptomology, and a clinically but not statistically significant benefit of lithium augmentation on the rate of remission, in adults with depression who have responded

Update 2017

inadequately to previous treatment. However, evidence from this study suggests a clinically important but not statistically significant harm associated with lithium augmentation relative to augmentation with low-intensity CBT as measured by discontinuation due to adverse events, although absolute numbers are small. Effects on discontinuation due to any reason were non-significant.

- Low quality single-study evidence (N=342) suggests a small but statistically significant benefit of augmenting antidepressant treatment with CBASP, relative to augmentation with short-term psychodynamic psychotherapy on depression symptomatology, and a trend for the same pattern of results on rate of remission, in adults with depression who have responded inadequately to previous treatment. However, very low quality evidence (N=395) from this same study suggests a clinically important but not statistically significant harm associated with CBASP augmentation relative to augmentation with short-term psychodynamic psychotherapy as measured by discontinuation due to adverse events, although absolute numbers are small. Effects on discontinuation due to any reason were non-significant.
- Very low quality evidence from 1-2 studies (N=29-102) suggests a clinically important but not statistically significant benefit of augmenting antidepressant treatment with exercise, relative to augmentation with attention-placebo, on the rate of remission, the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) and depression symptomatology, in adults with depression who have responded inadequately to previous treatment. However, very low quality evidence from these same 2 studies (N=106) suggests a clinically important but not statistically significant harm associated with exercise augmentation relative to attention-placebo as measured by discontinuation for any reason, although absolute numbers are relatively small.

8.5.3 Switching strategies

- Very low quality single-study evidence (N=322) suggests neither clinically important nor statistically significant benefits .of switching from an SSRI (paroxetine) to an atypical antidepressant (bupropion), relative to switching to placebo, on the rate of remission, the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or the number of participants rated as much or very much improved on the CGI-I), or on depression symptomatology in adults with depression who have responded inadequately to previous treatment. However, evidence from this study (N=325) did suggest a clinically important and statistically significant harm associated with switching to bupropion as measured by discontinuation for any reason, and a trend for the same pattern of results on discontinuation for adverse events
- Low to very low quality evidence from 3-4 studies (N=400-545) suggests neither clinically important nor statistically significant benefits of switching to an antidepressant of a different class, relative to continuing with the same antidepressant, on the rate of remission, the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or MADRS) or on depression symptomatology, in adults with depression who have responded inadequately to previous treatment. Evidence from all 4 studies (N=546-551) did, however, suggest a clinically important but not statistically significant harm associated with switching to an antidepressant of a different class as measured by discontinuation due to adverse events, and there was a trend for the same pattern of results with discontinuation due to any reason.
- Very low quality evidence from 2 studies (N=324-329) suggests neither clinically important nor statistically significant benefits of switching to an SSRI (fluoxetine), relative to continuing with the same TCA (nortriptyline) or SNRI (venlafaxine), on the rate of remission or response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS) or on depression symptomatology, in adults with depression who have responded inadequately to previous treatment. Evidence from

Update 2017

these 2 studies did, however, suggest a clinically important but not statistically significant harm associated with switching to an SSRI from an antidepressant of a different class as measured by discontinuation due to adverse events. No significant effects are shown on discontinuation due to any reason.

- Very low quality evidence from 2 studies (N=221) suggests neither a clinically important nor statistically significant benefit of switching to an atypical antidepressant (mirtazapine) or an SNRI (venlafaxine) or a TeCA (mianserin), relative to continuing with the same SSRI (fluoxetine or paroxetine), on the rate of remission or response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D) in adults with depression who have responded inadequately to previous treatment. Evidence from these 2 studies (N=217-222) did, however, suggest a clinically important but not statistically significant harm associated with switching to an antidepressant of a different class from an SSRI as measured by discontinuation for any reason and discontinuation due to adverse events.
- Very low quality single-study evidence (N=71) suggests clinically important but not statistically significant benefits of switching to a TeCA (mianserin), relative to continuing with the same SSRI (fluoxetine), on the rate of remission and response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or the number of participants rated as much or very much improved on the CGI-I) in adults with depression who have responded inadequately to previous treatment. However, this study found neither a clinically important nor statistically significant benefit of switching to mianserin on depression symptomatology. Evidence from this study suggests a clinically important and statistically significant harm associated with switching to mianserin from fluoxetine as measured by discontinuation due to adverse events, and a clinically important but not statistically significant harm as measured by discontinuation for any reason.
- Very low quality evidence from 3 studies (N=729) suggests a clinically important and statistically significant benefit in favour of continuing with the antidepressant relative to switching to antipsychotic monotherapy (olanzapine) on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS) in adults with depression who have responded inadequately to previous treatment. Evidence from the same 3 studies suggests a trend for the same pattern of results on the rate of remission and depression symptomatology. Evidence from these same 3 studies (N=738) also suggests a clinically important and statistically significant harm associated with switching to antipsychotic monotherapy as measured by discontinuation for any reason and discontinuation due to adverse events.
- Low to very low quality evidence from 2 studies (N=502-516) suggests neither clinically important nor statistically significant benefits of switching to combined antipsychotic and SSRI treatment (olanzapine + fluoxetine), relative to continuing with TCA (nortriptyline) or SNRI (venlafaxine) treatment, on the rate of remission, the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS), or depression symptomatology in adults with depression who have responded inadequately to previous treatment. Very low quality evidence from these same 2 studies does, however, suggest a clinically important and statistically significant harm associated with switching to combined antipsychotic and SSRI treatment as measured by discontinuation due to adverse events, and a trend for the same pattern of results with discontinuation due to any reason.
- Low quality evidence from 2 studies (N=460) suggests a clinically important and statistically significant benefit in favour of augmenting (with mianserin or quetiapine), relative to switching to (mianserin or quetiapine), on depression symptomatology in adults with depression who have responded inadequately to previous treatment.
- Low to very low quality single-study (N=65) suggests neither clinically important nor statistically significant benefits of switching to, relative to augmenting with, a TeCA (mianserin) on the rate of remission, the rate of response (as measured by the number of

Update 2017

participants showing at least 50% improvement from baseline on the HAM-D or the number of participants rated as much or very much improved on the CGI-I) or on depression symptomatology in adults with depression who have responded inadequately to previous treatment. Very low quality evidence from this same study (N=66) did, however, suggest a clinically important but not statistically significant harm associated with switching to (relative to augmenting with) mianserin as measured by discontinuation for any reason and discontinuation due to adverse events.

- Very low quality evidence from 2 studies (N=849) suggests a clinically important and statistically significant benefit in favour of augmenting with, relative to switching to, an antipsychotic on the rate of remission in adults with depression who have responded inadequately to previous treatment, and very low quality evidence from 1 of these studies (N=395) suggests a small but statistically significant benefit of augmenting relative to switching on depression symptomatology. However, evidence from these same studies (K=1-2; N=454-849) suggests neither clinically important nor statistically significant differences between antipsychotic augmentation and switching on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS or the number of participants rated as much or very much improved on the CGI-I). Very low quality evidence from these studies (N=858) suggests a clinically important and statistically significant harm associated with switching to (relative to augmenting with) an antipsychotic as measured by discontinuation for any reason, and a trend for the same pattern of results with discontinuation due to adverse events.
- Very low quality single-study evidence (N=446) suggests neither clinically important nor statistically significant benefits of switching to an antipsychotic, relative to augmenting usual antidepressant treatment with lithium, on the rate of remission, the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS or the number of participants rated as much or very much improved on the CGI-I) or discontinuation for any reason, in adults with depression who have responded inadequately to previous treatment. Evidence from this study (N=457) does, however, suggest a clinically important but not statistically significant harm associated with switching to an antipsychotic relative to lithium augmentation as measured by discontinuation for adverse events.
- Very low quality evidence from 2 studies (N=884) suggests a clinically important and statistically significant benefit of switching to an SNRI (different class), relative to switching to another SSRI (same class), on the rate of remission in adults with depression who have responded inadequately to previous treatment. However, low to very low quality evidence from 1 of these studies (N=488) suggests neither clinically important nor statistically significant differences between a same class (SSRI) and different class (SNRI) switch on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the QIDS) or depression symptomatology. Very low quality evidence from 1 or both of these studies (N=406-891) suggests neither clinically important nor statistically significant differences between same class and different class switch in terms of discontinuation for any reason or discontinuation due to adverse events.
- Low to very low quality single-study evidence (N=477) suggests neither clinically important nor statistically significant differences between switching to another antidepressant of the same class (SSRI), relative to switching to another antidepressant of a different class (atypical antidepressant), on the rate of remission or response (as measured by the number of participants showing at least 50% improvement from baseline on the QIDS), on depression symptomatology or on discontinuation due to adverse events, in adults with depression who have responded inadequately to previous treatment.
- Low quality evidence from 4 studies (N=1397) suggests a clinically important and statistically significant benefit of switching to a non-SSRI antidepressant, relative to an SSRI, on the rate of remission in adults with depression who have responded inadequately to previous treatment. However, very low quality evidence from 3 of these studies (N=718-1253) suggests neither clinically important nor statistically significant

Update 2017

benefits on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or QIDS) or depression symptomatology, or harms as measured by discontinuation for any reason or discontinuation due to adverse events.

- Very low quality evidence from 2 studies (N=401-408) suggests a clinically important and statistically significant benefit of switching to an SSRI, relative to an antipsychotic, on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS), and a small but statistically significant benefit on depression symptomatology, in adults with depression who have responded inadequately to previous treatment. However, evidence from the same 2 studies suggests neither a clinically important nor statistically significant benefit of switching to an SSRI, relative to switching to an antipsychotic, on the rate of remission or on discontinuation for any reason. Evidence from these same 2 studies suggests a clinically important and statistically significant harm associated with switching to an antipsychotic, relative to switching to an SSRI, as measured by discontinuation due to adverse events.
- Very low quality evidence from 2 studies (N=594) suggests neither a clinically important nor statistically significant benefit of switching to an SNRI, relative to an atypical antidepressant, on the rate of remission or the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the HAM-D or QIDS) in adults with depression who have responded inadequately to previous treatment. Very low quality evidence from 1-2 of these studies (N=105-589) also suggests neither clinically important nor statistically significant harms of switching to an SNRI relative to an atypical antidepressant, as measured by discontinuation for any reason or discontinuation due to adverse events.
- Very low quality evidence from 2 studies (N=579) suggests a clinically important but not statistically significant benefit of switching to a combined SSRI and antipsychotic, relative to switching to an antipsychotic-only, on the rate of remission and the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS) in adults with depression who have responded inadequately to previous treatment. Evidence from these 2 studies (N=595) suggests a trend for the same pattern of results on depression symptomatology and also suggests neither clinically nor statistically significant harms associated with switching to a combined SSRI and antipsychotic relative to switching to an antipsychotic-only, as measured by discontinuation for any reason and discontinuation due to adverse events.
- Very low quality evidence from 2 studies (N=574) suggests a clinically important but not statistically significant benefit of switching to a combined SSRI and antipsychotic, relative to switching to an SSRI-only, on the rate of remission in adults with depression who have responded inadequately to previous treatment. However, evidence from the same 2 studies (N=574-591) suggests neither a clinically important nor statistically significant benefit of switching to a combined SSRI and antipsychotic relative to switching to an SSRI-only on the rate of response (as measured by the number of participants showing at least 50% improvement from baseline on the MADRS), on depression symptomatology or on discontinuation due to any reason. Evidence from both these studies suggests a clinically important and statistically significant harm associated with switching to a combined SSRI and antipsychotic, relative to switching to an SSRI-only, as measured by discontinuation due to adverse events.
- Very low quality single-study evidence (N=107) suggests neither a clinically important nor statistically significant benefit of switching to an SNRI, relative to switching to an SSRI, on the rate of response (as measured by the number of participants rated as much or very much improved on the CGI-I) in adults with depression who have responded inadequately to previous treatment.
- Very low quality single-study evidence (N=22) suggests a clinically important but not statistically significant harm (as measured by discontinuation for any reason) of switching to a combined CBT under 15 sessions and antipsychotic intervention, relative to switching

Update 2017

to a CBT intervention-only, in adults with depression who have responded inadequately to previous treatment.

8.6 Economic evidence statements

8.6.1 Dose escalation strategies

- No economic evidence was identified

8.6.2 Augmentation strategies

- Evidence from 1 UK model-based study suggests that lithium dominates antipsychotics as an adjunct to SSRIs in the treatment of adults with treatment-resistant depression. The study is directly applicable to the NICE decision-making context and is characterised by potentially serious limitations.
- Evidence from 1 US model-based economic study suggests that augmentation strategies are more cost-effective than continuation of current antidepressant treatment in adults with major depression that failed to respond to previous treatment. The study is partially applicable to the UK context and is characterised by very serious limitations.
- Evidence from 1 US model-based economic study was inconclusive as to whether antipsychotics used as adjuncts to antidepressant therapy were cost-effective compared with antidepressant therapy alone in adults with major depression who had responded inadequately to previous antidepressant therapy, as the study did not use the QALY as the measure of outcome. The study is partially applicable to the UK context and is characterised by very serious limitations.

8.6.3 Switching strategies

- Evidence from 1 single UK study conducted alongside a RCT (N = 469) suggests that CBT is a cost-effective treatment option in people with depression who have responded inadequately to previous treatment. This evidence is directly applicable to the NICE decision-making context and is characterised by minor limitations.
- Evidence from 1 single UK study conducted alongside a RCT (N=158) is inconclusive regarding the cost effectiveness of cognitive therapy in people who have responded inadequately to previous treatment and have residual depressive symptoms, as the outcome measure was not the QALY and interpretation of the results depends on the willingness to pay in order to avoid an additional relapse. This evidence, although it was conducted in the UK, is only partially applicable to the NICE decision-making context (due to lack of QALY estimation) and it characterised by minor limitations.
- Evidence from 1 UK model-based economic study suggests that duloxetine is more cost-effective than venlafaxine and mirtazapine in people with depression who have responded inadequately to previous antidepressant treatment with SSRIs. The study is directly applicable to the UK context but is characterised by potentially serious limitations.
- Evidence from 1 Swedish model-based economic study suggests that escitalopram is more cost-effective than duloxetine and venlafaxine in adults with major depression treated in primary care, who had had a history of treatment with another antidepressant within the previous 6 months. The study is partially applicable to the UK context and is characterised by potentially serious limitations.
- Evidence from 1 US model-based economic study suggests that switching to another antidepressant is more cost-effective than continuation of current antidepressant treatment in adults with major depression that failed to respond to previous treatment. The study is partially applicable to the UK context and is characterised by very serious limitations.

Update 2017

- Evidence from 1 US model-based economic study suggests that paroxetine controlled release and sertraline are less cost-effective compared with other SSRIs in adults with major depression who failed to achieve remission with previous treatment with SSRIs. The study is partially applicable to the UK context and is characterised by very serious limitations.

8.76 From evidence to recommendations

8.7.17 Relative values of different outcomes

Critical outcomes were remission, response as measured by an agreed percentage improvement in symptoms and/or by a dichotomous rating of much or very much improved, relapse and measures of symptom improvement on validated scales. Attrition from treatment (for any reason and due to adverse events) was also considered an important outcome.

8.7.22 Trade-off between clinical benefits and harms

In developing recommendations for depression that has not responded or where there has been a limited response to treatment, the GC drew on their knowledge and experience that a significant number of people may not adhere to the prescribed treatment regimen and personal or social factors could have a significant impact on a person's response to treatment. They therefore agreed that a review of these factors should be considered before initiating any additional treatment options. They also agreed that increasing contact for those receiving medication should be recommended, as in their view the support provided by a prescriber could have additional benefits in terms of supporting the individual in dealing with their depressive symptoms and ensuring proper concordance with the medication regime.

The GC considered the short term and long-term harms associated with the side effects of medication including for the SSRIs drowsiness, nausea, insomnia, agitation, restlessness and sexual problems. For the TCAs additional concerns include the potential for cardiotoxicity and associated increased risk in overdose. For lithium there were concerns about renal toxicity and a potential impact on thyroid function. For the antipsychotics concerns with weight gain and hyperlipidaemia and raised blood glucose were also considered. The GC took these factors into consideration and in particular the increased burden of harms that may arise with the use of a combination of medications. In developing the recommendations, the GC were mindful of the negative consequences of prolonged depressive episodes including not only the impact on the mental health of the individual and their family but also on an individual's physical health (depression is associated with poorer physical health outcomes) and the impact on education and employment. The GC agreed that the benefits of improving the outcome of a depressive episode outweighed the potential harms.

When developing the recommendations for further line treatment, the GC considered a number of factors including the relative strength of the evidence, the preference that service users may have for medication or psychological interventions and the adverse effects of medication, in particular when combinations of medications are used. The GC were aware, from established data on response curves to antidepressant treatment, that most people who respond to pharmacological interventions will have started to do so 3 to 4 weeks after initiation of treatment. Response curves are similar for psychological interventions but response to psychological interventions may initially be slower than to medication with people typically responding to treatment by 4 to 6 weeks. In developing their recommendations, the GC consider two main scenarios; first where a person had not responded to initial medication and secondly where a person had not responded to initial psychological therapy. Where there was a limited or no response to an initial single treatment with medication the GC recommended that a combination with psychological

Update 2017

intervention (specifically CBT, BA or IPT (for people with interpersonal difficulties)) should be used. In developing this recommendation, the GC drew on the evidence for first line treatments particularly in more severe depression where combination treatment was more clinically and cost-effective than medication alone. For people who had not responded to an initial psychological therapy the GC recommended a combination with medication, either adding an SSRI (for example, sertraline or citalopram) or mirtazapine. In developing this recommendation, the GC again drew on the evidence for first line treatments particularly in more severe depression where combination treatment was more clinically and cost –effective than medication alone. The GC however recognised that some people would not wish to continue with medication and so, drawing on their expert knowledge and experience and the data on first line treatments developed a recommendation that a person should have the option of changing to a psychological therapy alone. On the same principles where a person would not wish to continue with a psychological treatment they should switch to medication alone.

The GC also considered that for people, who had had no or a limited response to an initial treatment with medication but who do not want a psychological intervention, then combined drug treatment is a possible option. Combinations with an antidepressant of a different class, antipsychotics (aripiprazole, risperidone, quetiapine, olanzapine) and lithium were all identified in the reviews undertaken for this guideline as effective (i.e. they resulted in improved rates of remission or response and in depressive symptoms) in the treatment of no or limited response to initial treatment and so the GC decided to recommend them. However, the GC were aware that combinations of medication can result in a significant increase in side effect burden and therefore recommended that people should be informed about this so that they can decide if this increased burden is acceptable to them. The GC were also aware that a number of prescribers, including GPs, would not feel competent to initiate such combination treatment and therefore also recommended that specialist advice or assessment be sought before starting a combined medication strategy, particularly when using an antipsychotic or lithium.

The GC were aware that currently, a common approach to a limited or non-response to pharmacological interventions is to either increase the dose or switch to an alternative medication. However, the GC noted that the evidence reviewed in this guideline did not provide significant support for either of these two strategies as being effective. The GC considered that in a number of the trials which were reviewed in this guideline, the absence of benefit for switching or increasing the dose was likely to be due to the fact that those who were maintained on the original medication also improved. The GC were however aware that some people would not want to try a psychological intervention nor be willing to accept the increased side effect burden of combined drug treatment. Given this, the GC agreed to make a recommendation for switching to another antidepressant or increasing the dose. However, the GC were concerned about the limited evidence for these strategies and so also recommended close monitoring and a review of the treatment strategy if there is no response after 2 to 4 weeks. They also recommended discussion of other treatment options should take place and consideration be given to referral for specialist advice.

In developing the recommendations for treatment resistant depression the GC considered both the clinical evidence reviewed and the cost-effectiveness studies, particularly those from the UK. The GC decided to recommend, based on the evidence, psychological treatments (including CBASP and CBT) which have been specially developed for treatment resistant depression. The GC were also aware of the need for long-term treatment with antidepressants for people who have treatment resistant depression and the consequent potential for adverse side effects. They therefore made recommendations about what to do if people were benefiting from the medication but were at risk of stopping it because of the burden of adverse effects.

Update 2017

8.7.31 Trade-off between net health benefits and resource use

The GC considered the high healthcare costs and outcomes to the person associated with treatment failure and treatment-resistant depression compared with depression that has responded to treatment, and expressed the view that successful treatment, as expressed by full response to treatment and eventual remission, would lead to the optimal outcome to the person but also considerable cost-savings to the healthcare system.

The GC considered the available economic evidence on treatments for people with depression who have responded inadequately to previous treatment. They noted that UK evidence suggests that CBT may be a cost-effective treatment option in this population. Regarding drugs, evidence from the UK suggests that duloxetine is more cost-effective than venlafaxine and mirtazapine in people with depression who responded inadequately to previous treatment with SSRIs, and evidence from Sweden suggests that escitalopram is more cost-effective than duloxetine and venlafaxine in people who responded inadequately to previous antidepressant treatment. Other evidence from the UK suggests that lithium dominates antipsychotics as an adjunct to SSRIs in the treatment of adults with treatment-resistant depression. The GC noted that economic evidence on psychological interventions is characterised by minor limitations, whereas evidence on pharmacological interventions is characterised by potentially serious limitations. Other available non-UK evidence was not considered as it was characterised by very serious limitations.

The GC acknowledged that the economic evidence in this area is sparse and has limitations, and decided to draw additional information from the economic analysis of treatments of a new depressive episode that was undertaken for the guideline. According to the guideline economic analysis, pharmacological treatment, group psychological therapies (such as group CBT) and other low-intensity psychological and physical interventions were the most cost-effective options for the treatment of new episodes of less severe depression in adults. On the other hand, for populations with more severe depression, the combination of CBT individual with an antidepressant was likely to be the most cost-effective option for the treatment of new episodes, followed by group CBT, behavioural therapies and SSRIs.

Update 2017

Considering the available economic evidence, the GC recommended the combination of medication and psychological treatment for people who have responded inadequately to medication alone or to psychological intervention alone, and the possibility of changing the components of combination therapy in people who are already on a combination of medication and a psychological therapy.

The GC acknowledged that increasing the frequency and duration of appointments to support people responding inadequately to initial pharmacological treatment has modest resource implications, which, nevertheless, are expected to lead to better outcomes for the person and also be fully or partially offset by cost-savings further down the pathway if they result in better adherence and monitoring and, eventually, in a satisfactory response to treatment.

The GC considered that offering an SSRI or mirtazapine to people whose symptoms have not adequately responded to an initial psychological intervention would have minor resource implications as the intervention cost of providing antidepressant treatment is overall lower than that of an individual psychological intervention. Moreover, the GC noted that switching from psychological therapy that led to inadequate response to a different type of treatment would potentially result in better outcomes for the person and, therefore, reduction in further care costs.

The GC considered that increasing the dose of a well-tolerated drug, switching between antidepressants within the same or different class, or adding an antidepressant to existing medication (for example, adding a SSRI or mirtazapine) would have negligible resource implications in terms of the drug acquisition cost, as these drugs are available in generic

form. Switching from a drug that is causing side effects to another drug of the same or different class may lead to cost-savings and better outcomes for the person, if the new drug is better tolerated.

The GC acknowledged the additional costs associated with provision of lithium or antipsychotics in specialist settings or after consultation with a specialist. These costs relate to specialist staff time but also to monitoring costs and costs associated with side effects of lithium and antipsychotics. The GC considered the available UK evidence according to which lithium dominates antipsychotics as an adjunct to SSRIs in the treatment of adults with treatment-resistant depression, but noted that this evidence is characterised by potentially serious limitations. Based on the above considerations, the GC recommended combining antidepressants with an antipsychotic or lithium in specialist settings, or after consultation with a specialist, as an option only to people who had had no or partial response to initial medication and who do not want to try psychological interventions. In this population, alternative effective treatment options are limited and the GC expressed the view that the benefits of augmenting treatment with lithium or antipsychotics are likely to outweigh costs associated with their provision to this group.

8.7.4 Quality of evidence

All the evidence reviewed was of low or very low quality, reflecting the low number of studies for each comparator, the small numbers in most trials and the imprecision of most of the results.

8.7.5 Other considerations

When reviewing the evidence for further line treatment the GC had originally decided to separately examine the evidence base for treatment resistant depression (usually defined as no or limited response to two adequate courses of an antidepressant) from no or limited response to treatment. However, after carefully reviewing the trial populations and the variation in the criteria used to identify both no or limited response and treatment resistance the GC came to the view that there were considerable similarities and overlaps between the two populations and therefore decided to use the same data sets for both questions to inform the development of recommendations for no or limited response.

Update 2017

8.8 Recommendations

76. If a person with depression has had no response or a limited response to initial treatment (within 3-4 weeks for antidepressant medication or 4-6 weeks for psychological therapy or combined medication and psychological therapy), assess:

- whether there are any personal or social factors that might explain why the treatment isn't working
- whether the person has not been adhering to the treatment plan, including any adverse effects of medication.

Work with the person to try and address any problems raised. [new 2017]

77. If a person has had no response or a limited response to initial treatment after assessing the issues in recommendation 76, provide more support by increasing the number and length of appointments. Also consider:

- changing to a combination of psychological therapy and medication if the person is on medication only, or

- changing to psychological therapy alone, if the person is on medication only and does not want to continue with medication or
- changing to a combination of 2 different classes of medication, in specialist settings, or after consulting a specialist, if the person is on medication only or a combination of medication and psychological therapy and does not want to continue with psychological therapy. [new 2017]

78. When changing treatment for a person with depression who has had no response or a limited response to initial medication, consider:

- combining the medication with a psychological therapy (CBT, BA, or IPT) **or**
- switching to a psychological therapy (CBT, BA, or IPT) if the person wants to stop taking medication. [new 2017]

79. If a person has had no response or a limited response to initial medication and does not want to try a psychological therapy, and wants to try a combination of medications, inform them of the likely increase in their side-effect burden (including risk of serotonin syndrome).

80. If a person wants to try a combination of medications and is willing to accept an increased side-effect burden, consider:

- adding an antidepressant of a different class to their initial medication, for example, an SSRI with mirtazapine, in specialist settings, or after consulting a specialist
- combining an antidepressant with an antipsychotic[q] or lithium in specialist settings, or after consulting a specialist. [new 2017]

Update 2017

81. When changing treatment for a person with depression who has had no response or a limited response to initial psychological therapy consider:

- combining the psychological therapy with an SSRI, for example, sertraline or citalopram, or mirtazapine, **or**
- switching to an SSRI, for example, sertraline or citalopram, or mirtazapine if the person wants to stop the psychological therapy [new 2017]

82. If a person has had no response or a limited response to initial medication and does not want a psychological therapy or a combination of medications, consider:

- continuing with the current medication, with extra support, close monitoring and an increased dose if the medication is well tolerated, **or**
- switching to a medicine of a different class[r], **or**
- switching to medication of the same class if there are problems with tolerability. [new 2017]

[q] At the time of consultation (July 2017) antipsychotics (with the exception of quetiapine and flupenthixol) 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 Prescribing guidance: prescribing unlicensed medicines for further information.

r Take into account there is limited evidence to support routine increases in dose of antidepressants or switching to a drug of a different class in people who have not responded to initial treatment.

83. If a person's symptoms do not respond to a dose increase or switching to another antidepressant after 2–4 weeks, review the need for care and treatment and consider consulting with, or referring the person to, a specialist service. [new 2017]

84. For people with depression whose symptoms have not adequately responded to a combination of medication and a psychological therapy after 12 weeks, consider:

- alternatives to combined treatment (see recommendation 87)
- switching to a different psychological therapy, such as cognitive behavioural analysis system of psychotherapy (CBASP), CBT or MBCT (see recommendation 86). [new 2017]

85. If a person finds that their antidepressant medication is helping them but they are having side effects, consider switching to another antidepressant with a different side effect profile. [new 2017]

Update 2017

9 Chronic depression

9.1 Introduction

Although depression is often viewed as a brief self-limiting disorder, convergent evidence from longitudinal studies indicates that many cases follow a chronic, unremitting course:

- 22 - 33% at 1-year follow-up (Keller et al. 1986, Rush et al. 2006)
- 21% at 2-year follow-up (Keller et al. 1984)
- 12% at 5-year follow-up (Keller et al. 1992)
- 7% at 10-year follow-up (Mueller et al. 1996)
- 6% at 15-year follow-up (Keller and Boland 1998).

This persistence of depression in adults is formally referred to as 'chronic depression' when it has continued beyond 2 years (APA 2000, WHO 1992); and although this convention is to some extent arbitrary it nevertheless provides an important reference for our current evidence base. Within the period of persistence, evidence indicates considerable variability in the nature of 'chronic depression', including: a persistent major depressive episode (clinical depression) that waxes and wanes without ever reaching the prior state of wellbeing (remission); a persistent depressed state that never quite fully meets criteria for a major depressive episode, taking a milder, chronic form called 'dysthymia'; or an alternating state of dysthymia and major depression (sometimes called 'double depression').

The most recent revision of psychiatric classification (DSM-V) now conflates the terms 'dysthymia' (a relatively mild depressed state, sub-syndromal for major depression but persistent over 2 years) and 'chronic depression' (non-remitting major depression) under the heading 'persistent depression' (300.4), although additional specifiers for 'pure dysthymic syndrome' and 'persistent major depressive episode' remain. In this chapter, the term chronic depression is used to include major depressive disorder that has lasted at least 2 years, dysthymia, double depression and recurrent depression with incomplete remission between episodes.

Update 2017

Studies have associated chronic depression with particularly high rates of hospitalisation, functional impairment and suicide (Arnow and Constantino 2003). There is also some indication of relatively early lifetime onset (Nanni, Uher and Danese 2012). Given that in any case major depression has a lifetime population risk of around 30% (Kessler et al. 2012), with typical onset by the early-mid 20s (Kessler and Bromet 2013) and associated economic costs that remain high throughout the working lifespan (largely related to lost productivity) (Kessler, Foster and Saunders 1995, Whiteford et al. 2010), the absolute human and economic costs of its chronic form are likely to be substantial.

Given all of this, it may not be surprising that once depression has become chronic the outcome tends to be poor (Buszewicz et al. 2016). And yet, despite evidence on the persistence, cost, complexity and poor prognosis of chronic depression, research on treatment is both scarce (in comparison to early stage depression) and generally limited to single interventions (such as pharmacotherapy or psychotherapy) with few trials of combination (Keller et al. 2000) or service level, multi-professional interventions (Buszewicz et al. 2016, Murray et al. 2010). This chapter will assess this evidence base and the gaps within it.

9.2 Review question

- For adults with chronic depression what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 144. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 144: Clinical review protocol summary for the review of chronic depression

Topic	Treatment of chronic depression
Review question	For adults with chronic depression what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination? (RQ2.6)
Population	Adults with chronic depression, defined by a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on scales The definition of chronic depression includes: meeting criteria for full MDD for 2 years; persistent subthreshold symptoms (dysthymia); double depression (an acute episode of MDD superimposed on dysthymia) In the case of mixed populations, if the study reports data for a subgroup with chronic depression, data for this subgroup will be extracted. If the study does not report data separately we will only include studies where over 75% of the population have a diagnosis of chronic depression. Studies with mixed populations where less than 75% of the population have chronic depression will be included in other reviews.
Intervention	Interventions listed below are examples of interventions which may be included either alone or in combination. **Psychological interventions:** • Cognitive and cognitive behavioural therapies (including CBT individual or group [defined as under or over 15 sessions], cognitive behavioural analysis system of psychotherapy (CBASP), mindfulness-based cognitive therapy (MBCT), and problem solving) • Counselling • Interpersonal psychotherapy (IPT) • Psychodynamic psychotherapies **Psychosocial interventions**: • Befriending • Mentoring • Peer support • Community navigators **Pharmacological interventions**: • Antidepressants ○ SSRIs - citalopram - escitalopram - fluvoxamine - fluoxetine - paroxetine - sertraline ○ TCAs - amineptine[1] - amitriptyline - clomipramine - desipramine[2] - imipramine

Update 2017

Topic	Treatment of chronic depression
	- lofepramine - nortriptyline ○ MAOIs - phenelzine ○ TeCAs - mianserin ○ SNRIs - duloxetine - venlafaxine ○ Other antidepressant drugs - bupropion[3] - mirtazepine - moclobemide - nefazodone[2] • Antipsychotics ○ amisulpride[3] ○ aripiprazole[3] ○ olanzapine[3] ○ quetiapine[4] ○ risperidone[3] ○ ziprasidone[2] **Physical interventions**: • Acupuncture • ECT • Exercise (including yoga)
Comparison	• Treatment as usual • Waitlist • Placebo • Any other active comparison
Critical outcomes	• **Efficacy** • Depression symptomology (mean endpoint score or change in depression score from baseline) • Remission (usually defined as a cut off on a depression scale) • Response (e.g. reduction of at least 50% from the baseline score on HAMD/MADRS) • Relapse The following depression scales will be included in the following hierarchy: • MADRS • HAMD • QIDS • PHQ • CGI • CES-D • BDI • HADS-D (depression subscale) • HADS (full scale) • Acceptability/tolerability • Discontinuation due to any reason (including adverse events) • Discontinuation due to adverse events

Update 2017

Topic	Treatment of chronic depression
Study design	• RCTs • Cluster RCTs

Note: [1]Amineptine is not available to prescribe as a medicine (although it falls under Class C of the Misuse of Drugs Act 1971, and listed as Schedule 2 under the Controlled Drugs Regulations 2001). However, this drug is included in this review in order to assess the class effect of pharmacological interventions for depression
[2]These drugs are not available in the UK to prescribe. However, they are included in this review in order to assess the class effect of pharmacological interventions for depression
[3]None of these drugs are licensed for use in depression. However, they are included in the review in order to assess harms and efficacy for off-label use and to assess the class effect of pharmacological interventions for depression
[4]Quetiapine is licensed for use as an adjunctive treatment of major depressive episodes with major depressive disorder but not as monotherapy

9.3 Clinical evidence

Eighty-seven studies of treatment for chronic depression in adults were identified for full-text review. Of these 87 studies, 32 RCTs were included (Anisman 1999; Bakish 1993a; Barnhofer 2009; Boyer 1996 (study 1); Boyer 1996 (study 2)/Lecrubier 1997; Browne 2002; de Mello 2001; Dunner 1996; Hellerstein 1993; Hellerstein 2001; Hellerstein 2010; Hellerstein 2012; Keller 1998a; Keller 2000; Klein 2004; Kocsis 1988a; Markowitz 2005; Michalak 2015; Ravindran 2013; Schramm 2008; Schramm 2011; Schramm 2015; Schramm 2017; Stewart 1989/1993; Strauss 2012; Thase 1996; Vallejo 1987; Vanelle 1997; Versiani 1997; Wiersma 2014; Williams 2000; Wong 2008). Fifty-five studies were reviewed at full-text and excluded from this review. The most common reasons for exclusion were a non-chronic population (<80% of sample had depression for at least 2 years) or that the study included a mixed population, for instance, different diagnoses or chronic and non-chronic depression, and less than 80% of the sample met the inclusion criteria and it was not possible to extract disaggregated data. Studies not included in this review with reasons for their exclusions are provided in Appendix J5.

Update 2017

9.3.1 Psychological interventions for chronic depression

Evidence was found for five comparisons of cognitive and cognitive behavioural therapies as follows: cognitive and cognitive behavioural therapies compared to placebo (see Table 145 for study characteristics); cognitive and cognitive behavioural therapies compared to antidepressants (see Table 147 for study characteristics); cognitive and cognitive behavioural therapies compared to other psychological interventions (see Table 149 for study characteristics); cognitive and cognitive behavioural therapies in combination with antidepressants or treatment as usual compared to antidepressants or treatment as usual only (see Table 151 for study characteristics); cognitive and cognitive behavioural therapies compared to attention-placebo for relapse prevention (see Table 153 for study characteristics).

Evidence was found for three comparisons of IPT as follows: IPT compared to antidepressants (see Table 155 for study characteristics); IPT compared to other psychological interventions (see Table 157 for study characteristics); IPT in combination with antidepressants or treatment as usual compared to antidepressants or treatment as usual only (see Table 159 for study characteristics).

Evidence was found for two other psychological intervention comparisons as follows: brief supportive psychotherapy (BSP) compared to antidepressants (see Table 161 for study characteristics); Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD) combined with antidepressants compared to maintenance treatment with antidepressants-only for relapse prevention (see Table 163 for study characteristics).

Evidence for these comparisons are summarised in the clinical GRADE evidence profiles below (Table 146, Table 148, Table 150, Table 152, Table 154, Table 156, Table 158, Table 160, Table 162 and Table 164). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J6.

Table 145: Study information table for trials included in the meta-analysis of cognitive and cognitive behavioural therapies versus placebo

	Problem solving versus pill placebo
Total no. of studies (N randomised)	1 (142)
Study ID	Williams 2000
Country	US
Chronic definition	DSM-III-R dysthymia (confirmed with PRIME-MD; trial also included minor depression but data only extracted for subgroup with dysthymia)
Age range (mean)	NR
Sex (% female)	NR
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	NR
Intervention details	Problem-Solving Treatment-Primary Care (PST-PC; followed method of Mynors-Wallis 1996)
Intervention dose	6 sessions (1 hour for first session and 30-min subsequently)
Comparator details (mean dose, if applicable)	Pill placebo 10-40mg/day + clinical management (6x 15-min sessions of medication management)
Treatment length (weeks)	11
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation Williams 2000 is a three-armed trial but data extracted for the two relevant arms here, data also only extracted for dysthymia subgroup from this study and as a result demographic details limited (not reported by diagnostic subgroup)	

Table 146: Summary of findings table for cognitive and cognitive behavioural therapies 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				
	Pill placebo	Problem solving				
Remission Number of people scoring <7 on Hamilton Rating Scale for	**Study population**		**RR 1.26** (0.85 to 1.86)	125 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	403 per 1000	**508 per 1000** (343 to 750)				

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Pill placebo	Problem solving				
Depression (HAM-D) Follow-up: mean 11 weeks	**Moderate**					
	403 per 1000	**508 per 1000** (343 to 750)				

Notes:
[1] Intervention administrators and participants not blinded, although outcome assessment is blinded
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Medication and placebo supplied by pharmaceutical company and authors have some financial interests in pharmaceutical companies

Table 147: Study information table for trials included in the meta-analysis of cognitive and cognitive behavioural therapies versus antidepressants

	Cognitive and cognitive behavioural therapies versus antidepressants
Total no. of studies (N randomised)	4 (912)
Study ID	Dunner 19961 Keller 20002 Schramm 20153 Williams 20004
Country	US[1,2,4] Germany[3]
Chronic definition	Dysthymia[1,4] Mixed (35% MDD ≥2 years; 42% double depression; 23% recurrent depression with incomplete remission between episodes) [2] Double depression (63%; + 20% recurrent major depressive episodes [≥ 3 episodes with the preceding episode no more than 2.5 years before the onset of the current episode] and 14% MDD ≥1 year) [3]
Age range (mean)	19-50 (35.7) [1] Range NR (43) [2] Range NR (43.6) [3] NR[4]
Sex (% female)	46[1] 65[2] 54[3] NR[4]
Ethnicity (% BME)	NR[1,3,4] 9[2]
Mean age (SD) at first onset of depression	NR[1,3,4] MDD: 26.7 (13). Dysthymia: 19.3 (14) [2]
Mean months (SD) since onset of current episode	200 (134.8) [1] MDD: 93.6 (115.2). Dysthymia: 276 (180) [2] NR[3,4]

Update 2017

	Cognitive and cognitive behavioural therapies versus antidepressants
No. (SD) of previous depressive episodes	NR
Previous treatment	NR[1,4] 65% psychotherapy; 60% antidepressants; 45% both antidepressants and psychotherapy; 20% no prior treatment for depression[2] 68% psychotherapy; 60% medication; 47% both; 24% neither type of treatment. 21% at least 2 self-reported failures/nonresponses to a medication; 9% treatment-resistant to a psychotherapy course of at least 10 sessions[3]
Baseline severity	HAMD 16 (Less severe) [1] HAMD 26.9 (More severe) [2] MADRS 26.2 (Less severe) [3] NR[4]
Intervention details	CBT (followed the manual by Beck et al. 1979) [1] CBASP (followed the manual by McCullough 1995) [2] CBASP (followed the manual by McCullough 2000; German version: Schramm et al. 2006) [3] Problem-Solving Treatment-Primary Care (PST-PC; followed method of Mynors-Wallis 1996) [4]
Intervention dose	16x weekly sessions[1] 16-20 sessions (mean attended 16.0 sessions [SD=4.7]) [2] 12 sessions[3] 6 sessions (1 hour for first session and 30-min subsequently) [4]
Comparator details (mean dose, if applicable)	Fluoxetine 20mg/day + clinical management (weekly/biweekly 15-20 min sessions on medication management) [1] Nefazodone 200-600mg/day (final mean dose 466mg [SD=144]) [2] Escitalopram 10-20mg/day + clinical management (8x weekly 20-min sessions of clinical management) [3] Paroxetine 10-40mg/day + clinical management (6x 15-min sessions of medication management) [4]
Treatment length (weeks)	16[1] 12[2] 8[3] 11[4]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Dunner 1996; [2]Keller 2000; [3]Schramm 2015; [4]Williams 2000 Williams 2000[4] is a three-armed trial but data extracted for the two relevant arms here, data also only extracted for dysthymia subgroup from this study and as a result demographic details limited (not reported by diagnostic subgroup)	

Update 2017

Table 148: Summary of findings table for cognitive and cognitive behavioural therapies versus antidepressants

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Antidepressants	Cognitive and cognitive behavioural therapies				
Remission (any cognitive or cognitive behavioural therapy versus any AD) Number of people scoring <7/≤8 on Hamilton Rating Scale for Depression (HAM-D)/ ≤9 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 1.1** (0.83 to 1.46)	615 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	309 per 1000	**340 per 1000** (257 to 452)				
	Moderate					
	291 per 1000	**320 per 1000** (242 to 425)				
Remission (CBASP versus nefazodone) Number of people scoring ≤8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 12 weeks	**Study population**		**RR 1.15** (0.87 to 1.52)	436 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	291 per 1000	**335 per 1000** (253 to 442)				
	Moderate					
	291 per 1000	**335 per 1000** (253 to 442)				
Remission (CBASP versus escitalopram) Number of people scoring ≤9 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 8 weeks	**Study population**		**RR 0.21** (0.03 to 1.67)	59 (1 study)	⊕⊖⊖⊖ **very low**[4,5,6]	
	167 per 1000	**35 per 1000** (5 to 278)				
	Moderate					
	167 per 1000	**35 per 1000** (5 to 279)				
Remission (problem solving versus paroxetine) Number of people scoring <7 on Hamilton Rating	**Study population**		**RR 1.11** (0.77 to 1.62)	120 (1 study)	⊕⊖⊖⊖ **very low**[5,6,7]	
	456 per 1000	**506 per 1000** (351 to 739)				
	Moderate					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Antidepressants	Cognitive and cognitive behavioural therapies				
Scale for Depression (HAM-D) Follow-up: mean 11 weeks	**456 per 1000**	**506 per 1000** (351 to 739)				
Response (any cognitive or cognitive behavioural therapy versus any AD) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND HAMD score 8-15)/≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 8-12 weeks	**Study population**		**RR 0.56** (0.21 to 1.49)	495 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,5,8]	
	196 per 1000	**110 per 1000** (41 to 292)				
	Moderate					
	227 per 1000	**127 per 1000** (48 to 338)				
Response (CBASP versus nefazodone) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND HAMD score 8-15 Follow-up: mean 12 weeks	**Study population**		**RR 0.77** (0.5 to 1.18)	436 (1 study)	⊕⊖⊖⊖ **very low**[1,3,9]	
	186 per 1000	**144 per 1000** (93 to 220)				
	Moderate					
	186 per 1000	**143 per 1000** (93 to 219)				
Response (CBASP versus escitalopram) Number of people showing ≥50% improvement on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 8 weeks	**Study population**		**RR 0.26** (0.06 to 1.12)	59 (1 study)	⊕⊖⊖⊖ **very low**[6,7,9]	
	267 per 1000	**69 per 1000** (16 to 299)				
	Moderate					
	267 per 1000	**69 per 1000** (16 to 299)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Antidepressants	Cognitive and cognitive behavioural therapies				
Depression symptomatology (any cognitive or cognitive behavioural therapy versus any AD) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: 12-16 weeks		The mean depression symptomatology (any cognitive or cognitive behavioural therapy versus any AD) in the intervention groups was **0.61 standard deviations higher** (0.54 lower to 1.76 higher)		458 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,10,11]	SMD 0.61 (-0.54 to 1.76)
Depression symptomatology (CBASP versus nefazodone) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 12 weeks		The mean depression symptomatology (CBASP versus nefazodone) in the intervention groups was **0.11 standard deviations higher** (0.08 lower to 0.3 higher)		436 (1 study)	⊕⊕⊖⊖ **low**[1,3]	SMD 0.11 (-0.08 to 0.3)
Depression symptomatology (CBT versus fluoxetine) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 16 weeks		The mean depression symptomatology (CBT versus fluoxetine) in the intervention groups was **1.3 standard deviations higher** (0.36 to 2.24 higher)		22 (1 study)	⊕⊕⊖⊖ **low**[12,13]	SMD 1.3 (0.36 to 2.24)
Discontinuation for any reason (any cognitive or cognitive behavioural therapy versus any	**Study population**		**RR 0.92** (0.68 to 1.25)	545 (3 studies)	⊕⊖⊖⊖ **very low**[3,7,14]	
	248 per 1000	**228 per 1000** (169 to 310)				
	Moderate					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Antidepressants	Cognitive and cognitive behavioural therapies				
AD) Number of participants discontinuing for any reason including adverse events Follow-up: 8-16 weeks	**231 per 1000**	**213 per 1000** (157 to 289)				
Discontinuation for any reason (CBASP versus nefazodone) Number of participants discontinuing for any reason including adverse events Follow-up: mean 12 weeks	**Study population**		**RR 0.92** (0.67 to 1.27)	454 (1 study)	⊕⊖⊖⊖ **very low**[3,7,14]	
	261 per 1000	**240 per 1000** (175 to 332)				
	Moderate					
	261 per 1000	**240 per 1000** (175 to 331)				
Discontinuation for any reason (CBASP versus escitalopram) Number of participants discontinuing for any reason including adverse events Follow-up: mean 8 weeks	**Study population**		**RR 0.43** (0.09 to 2.03)	60 (1 study)	⊕⊖⊖⊖ **very low**[4,6,14]	
	161 per 1000	**69 per 1000** (15 to 327)				
	Moderate					
	161 per 1000	**69 per 1000** (14 to 327)				
Discontinuation for any reason (CBT versus fluoxetine) Number of participants discontinuing for any reason including adverse events Follow-up: mean 16 weeks	**Study population**		**RR 1.44** (0.44 to 4.74)	31 (1 study)	⊕⊖⊖⊖ **very low**[14,15]	
	231 per 1000	**332 per 1000** (102 to 1000)				
	Moderate					
	231 per 1000	**333 per 1000** (102 to 1000)				
Discontinuation due to adverse events (CBASP versus nefazodone) Number of	**Study population**		**RR 0.1** (0.03 to 0.31)	454 (1 study)	⊕⊖⊖⊖ **very low**[3,7,16]	
	137 per 1000	**14 per 1000** (4 to 43)				
	Moderate					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Antidepressants	Cognitive and cognitive behavioural therapies				
participants discontinuing due to adverse events Follow-up: mean 12 weeks	**137 per 1000**	**14 per 1000** (4 to 42)				

Notes:
[1] Non-blind intervention administrator(s) and participants, although the outcome assessor was blinded. Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Funding from pharmaceutical company
[4] Unclear method of allocation concealment and non-blind intervention administrator(s) and participants, although the outcome assessor was blinded
[5] 95% CI crosses line of no effect and both threshold for clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[6] Data cannot be extracted or is not reported for all outcomes and funding from pharmaceutical company
[7] Non-blind intervention administrator(s) and participants, although outcome assessors are blinded
[8] I-squared=>50%
[9] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 0.75)
[10] I-squared>80%
[11] 95% CI crosses line of no effect and threshold for both clinically important benefit (SMD -0.5) and clinically important harm (SMD 0.5)
[12] Unclear randomisation method and method of allocation concealment, and non-blind intervention administrator(s) and participants (although outcome assessors are blinded). Unclear risk of attrition bias (drop-out>20% and completer analysis used but difference between groups<20%)
[13] N<400
[14] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[15] Unclear randomisation method and method of allocation concealment, and non-blind intervention administrator(s) and participants (although outcome assessors are blinded)
[16] Events<300

Table 149: Study information table for trials included in the meta-analysis of cognitive and cognitive behavioural therapies versus other psychological interventions

	CBASP versus other psychological intervention
Total no. of studies (N randomised)	2 (298)
Study ID	Schramm 2011[1] Schramm 2017[2]
Country	Germany
Chronic definition	Double depression (55%; + 31% early-onset [<21 years old] chronic MDD and 13% dysthymia) [1] Mixed (32% early-onset [<21 years old] chronic MDD, 46% double depression and 23%

Update 2017

	CBASP versus other psychological intervention
	recurrent major depressive disorder with incomplete inter-episode recovery) [2]
Age range (mean)	20-60 (40.2) [1] Range NR (44.9) [2]
Sex (% female)	55[1] 66[2]
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR[1] 13.0 (4.4) [2]
Mean months (SD) since onset of current episode	243.6 (135.6) [1] NR[2]
No. (SD) of previous depressive episodes	NR
Previous treatment	72% psychotherapy; 59% pharmacotherapy; 21% no prior treatment. 45% indicated no response to at least 2 previous trials of psychotherapy, 41% reported treatment resistance to antidepressants, 24% of those were resistant to both medication and psychotherapy trials[1] 57% psychotherapy; 55% medication; 20% combination. 52% previous inpatient treatment. Treatment-resistance (2 self-reported failures or nonresponses to a medication [>4 weeks] or to psychotherapy [>8 sessions]): 10% medication and 10% psychotherapy[2]
Baseline severity	HAMD 23.2 (Less severe) [1] HAMD 24.7 (More severe) [2]
Intervention details	CBASP (followed the manual by McCullough 2000; German version: Schramm et al. 2006)
Intervention dose	22-24x once/twice weekly 50-min sessions (mean attended 21.21 sessions [SD=3.12]) [1] 24 sessions[2]
Comparator details (mean dose, if applicable)	IPT (followed the manual by Klerman et al. 1984 and Weissman et al. 2000; German version: Schramm 1998; and modified for use with chronic depression by Markowitz 1998). 22-24x once/twice weekly 50-min sessions (mean attended 21.21 sessions [SD=3.12]) [1] Supportive psychotherapy (Markowitz 2014). 24 sessions[2]
Treatment length (weeks)	16[1] 20[2]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Schramm 2011; [2]Schramm 2017	

Update 2017

Table 150: Summary of findings table for cognitive and cognitive behavioural therapies versus other psychological 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				
	Other psych intervention	CBASP				
Remission (CBASP versus other psych intervention) Number of people scoring ≤8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 16-20 weeks	**Study population**		**RR 1.93** (1.14 to 3.26)	264 (2 studies)	⊕⊖⊖⊖ **very low**[1,2]	
	135 per 1000	**260 per 1000** (154 to 440)				
	Moderate					
	163 per 1000	**315 per 1000** (186 to 531)				
Remission (CBASP versus IPT) Number of people scoring ≤8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 16 weeks	**Study population**		**RR 2.86** (0.94 to 8.66)	29 (1 study)	⊕⊕⊖⊖ **low**[3,4]	
	200 per 1000	**572 per 1000** (188 to 1000)				
	Moderate					
	200 per 1000	**572 per 1000** (188 to 1000)				
Remission (CBASP versus supportive psychotherapy) Number of people scoring ≤8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 20 weeks	**Study population**		**RR 1.73** (0.95 to 3.12)	235 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	
	126 per 1000	**218 per 1000** (120 to 394)				
	Moderate					
	126 per 1000	**218 per 1000** (120 to 393)				
Response (CBASP versus other psych intervention) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 16-20 weeks	**Study population**		**RR 1.7** (1.18 to 2.44)	264 (2 studies)	⊕⊖⊖⊖ **very low**[1,2]	
	246 per 1000	**418 per 1000** (290 to 600)				
	Moderate					
	255 per 1000	**434 per 1000** (301 to 622)				
Response (CBASP versus IPT) Number of people showing ≥50%	**Study population**		**RR 2.41** (0.96 to 6.08)	29 (1 study)	⊕⊕⊖⊖ **low**[3,4]	
	267 per 1000	**643 per 1000** (256 to 1000)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Other psych intervention	CBASP				
improvement on Hamilton Rating Scale for Depression (HAM-D) AND HAMD score≤15) Follow-up: mean 16 weeks	**Moderate**					
	267 per 1000	**643 per 1000** (256 to 1000)				
Response (CBASP versus supportive psychotherapy) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 20 weeks	**Study population**		**RR 1.59** (1.07 to 2.36)	235 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	243 per 1000	**387 per 1000** (260 to 574)				
	Moderate					
	243 per 1000	**386 per 1000** (260 to 573)				
Depression symptomatology (CBASP versus other psych intervention) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: 16-20 weeks		The mean depression symptomatology (CBASP versus other psych intervention) in the intervention groups was **0.49 standard deviations lower** (0.98 lower to 0 higher)		297 (2 studies)	⊕⊖⊖⊖ **very low**[1,5]	SMD -0.49 (-0.98 to 0)
Depression symptomatology (CBASP versus IPT) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 16 weeks		The mean depression symptomatology (CBASP versus IPT) in the intervention groups was **0.89 standard deviations lower** (1.66 to 0.12 lower)		29 (1 study)	⊕⊕⊖⊖ **low**[3,5]	SMD -0.89 (-1.66 to -0.12)
Depression symptomatology (CBASP versus supportive psychotherapy) Hamilton Rating Scale for Depression (HAM-D; change score)		The mean depression symptomatology (CBASP versus supportive psychotherapy) in the intervention groups was **0.33 standard**		268 (1 study)	⊕⊖⊖⊖ **very low**[1,5]	SMD -0.33 (-0.58 to -0.09)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Other psych intervention	CBASP				
Follow-up: mean 20 weeks		**deviations lower** (0.58 to 0.09 lower)				
Discontinuation for any reason (CBASP versus other psych intervention) Number of participants discontinuing for any reason including adverse events Follow-up: 16-20 weeks	**Study population**		**RR 0.64** (0.35 to 1.16)	298 (2 studies)	⊕⊖⊖⊖ **very low**[1,6]	
	164 per 1000	**105 per 1000** (58 to 191)				
	Moderate					
	151 per 1000	**97 per 1000** (53 to 175)				
Discontinuation for any reason (CBASP versus IPT) Number of participants discontinuing for any reason including adverse events Follow-up: mean 16 weeks	**Study population**		**RR 1** (0.16 to 6.2)	30 (1 study)	⊕⊖⊖⊖ **very low**[3,7]	
	133 per 1000	**133 per 1000** (21 to 827)				
	Moderate					
	133 per 1000	**133 per 1000** (21 to 825)				
Discontinuation for any reason (CBASP versus supportive psychotherapy) Number of participants discontinuing for any reason including adverse events Follow-up: mean 20 weeks	**Study population**		**RR 0.61** (0.33 to 1.14)	268 (1 study)	⊕⊖⊖⊖ **very low**[1,6]	
	168 per 1000	**102 per 1000** (55 to 191)				
	Moderate					
	168 per 1000	**102 per 1000** (55 to 192)				

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded
[2] Events<300
[3] Non-blind intervention administrator(s) and participants, although outcome assessors are blinded
[4] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[5] N<400
[6] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 0.75)
[7] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Update 2017

Table 151: Study information table for trials included in the meta-analysis of cognitive and cognitive behavioural therapies in combination with antidepressants or treatment as usual versus antidepressants or treatment as usual-only

	Cognitive and cognitive behavioural therapies + TAU/AD versus TAU/AD-only
Total no. of studies (N randomised)	6 (1084)
Study ID	Barnhofer 2009[1] Keller 2000[2] Michalak 2015[3] Strauss 2012[4] Wiersma 2014[5] Wong 2008[6]
Country	UK[1,4] US[2] Germany[3] Netherlands[5] China[6]
Chronic definition	MDD ≥2 years (75%; + 25% residual symptoms following a full episode) [1] Mixed (35% MDD ≥2 years; 42% double depression; 23% recurrent depression with incomplete remission between episodes) [2] MDD ≥2 years (83%) [3] Unclear (DSM-IV chronic major depression, recurrent depression without full inter-episode recovery or double depression) [4,5] Unclear (DSM-IV MDD [mean duration of illness = 5.5 years]) [6]
Age range (mean)	Range NR (41.9) [1] Range NR (43) [2] Range NR (50.8) [3] Range NR (43) [4] Range NR (41.6) [5] Range NR (37.4) [6]
Sex (% female)	68[1] 65[2] 62[3] 71[4] 60[5] 78[6]
Ethnicity (% BME)	NR[1,3,4,5,6] 9[2]
Mean age (SD) at first onset of depression	21.9 (9.8) [1] MDD: 26.7 (13). Dysthymia: 19.3 (14) [2] NR[3,6] 20 (8) [4] 24.4 (12.8) [5]
Mean months (SD) since onset of current episode	101.9 (103.6) [1] MDD: 93.6 (115.2). Dysthymia: 276 (180) [2] NR[3,5] 48 (range 24-120) [4] 66 (57.6) [6]

Update 2017

	Cognitive and cognitive behavioural therapies + TAU/AD versus TAU/AD-only
No. (SD) of previous depressive episodes	5.4 (9.4) [1] NR[2,3,4,5] 2.6 (SD NR) [6]
Previous treatment	75% psychotherapy or counselling; 54% CBT; 82% antidepressants[1] 65% psychotherapy; 60% antidepressants; 45% both antidepressants and psychotherapy; 20% no prior treatment for depression[2] NR[3,6] 84% psychotherapy[4] 82% previous mental health treatment (secondary or tertiary care) [5]
Baseline severity	BDI-II 30.3 (More severe) [1] HAMD 26.9 (More severe) [2] HAMD 23.9 (Less severe) [3] BDI-II 39.1 (More severe) [4] IDS 42.4 (Unclear) [5] BDI 23.9 (More severe) [6]
Intervention details	MBCT (followed the manual by Segal et al. 2002) + TAU (14% changed antidepressant medication; 29% received psychological intervention; 57% visited GP regarding depression; 29% received visit by psychiatric nurse; 43% use of self-help [books etc.]) [1] CBASP (followed the manual by McCullough 1995) + nefazodone[2] MBCT (followed the manual by Segal et al. 2002) + TAU (75% receiving antidepressants and 32% individual psychotherapy) [3] CBASP (followed the manual by McCullough 2000, and modified for the group setting by Schramm et al. 2012)+ TAU (76% receiving antidepressants and 40% individual psychotherapy) [3] Person-Based Cognitive Therapy (PBCT) (modified version of Chadwick 2006 and Dannahy et al. 2011) + TAU (88% on antidepressant medication) [4] CBASP (followed the manual by McCullough 1995) + TAU[5] CBT group (followed manual by Greenberger & Padesky 1995) + TAU (all participants taking medication, almost all taking TCAs or SSRIs)[6]
Intervention dose	8x weekly 2-hour sessions (mean attended 6.14 sessions [SD=1.51]) [1] 16-20 sessions (mean attended 16.2 sessions [SD=4.8]) + 200-600mg/day of nefazodone (mean final dose 460mg [SD=139]) [2] 8x weekly 2.5-hour sessions[3] 12x weekly 90-min sessions (mean attended 8.92 sessions [SD=3.57]) [4] 24x 45-min sessions (mean attended 24.3 sessions [SD=10.8]) [5]

Update 2017

	Cognitive and cognitive behavioural therapies + TAU/AD versus TAU/AD-only
	10x weekly 2.5 hour sessions[6]
Comparator details (mean dose, if applicable)	TAU (50% changed antidepressant medication; 43% received psychological intervention; 50% visited GP regarding depression; 29% received visit by psychiatric nurse; 43% use of self-help [books etc.]) [1] nefazodone 200-600mg/day (final mean dose 466mg [SD=144]) [2] TAU (53% receiving antidepressants and 38% individual psychotherapy) [3] TAU (88% on antidepressant medication) [4] TAU (95% psychotherapy [53% CBT; 25% IPT; 10% short-term psychodynamic psychotherapy; 7% supportive/structured therapy]; 60% antidepressant use; 5% pharmacotherapy only) [5] Waitlist + TAU (all participants taking medication, almost all taking TCAs or SSRIs) [6]
Treatment length (weeks)	8[1,3] 12[2,4] 52[5] 10[6]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Barnhofer 2009; [2]Keller 2000; [3]Michalak 2015; [4]Strauss 2012; [5]Wiersma 2014; [6]Wong 2008	

Table 152: Summary of findings table for cognitive and cognitive behavioural therapies in combination with antidepressants or treatment as usual compared to antidepressants or treatment as usual only

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	Cognitive and cognitive behavioural therapies + TAU/AD				
Remission (MBCT+TAU versus TAU) Number of people scoring ≤13 on Beck Depression Inventory II (BDI-II) AND ≥50% improvement on BDI-II/≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 3.72** (1.1 to 12.54)	102 (2 studies)	⊕⊖⊖⊖ **very low**[1,2]	
	60 per 1000	**223 per 1000** (66 to 752)				
	Moderate					
	62 per 1000	**231 per 1000** (68 to 777)				
	Study population					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	Cognitive and cognitive behavioural therapies + TAU/AD				
Remission (CBASP + TAU/nefazodone versus TAU/nefazodone) Number of people scoring ≤7/8 on Hamilton Rating Scale for Depression (HAM-D)/≤13 on Inventory of Depressive Symptoms (IDS) Follow-up: 8-52 weeks	**227 per 1000**	**388 per 1000** (306 to 488)	**RR 1.71** (1.35 to 2.15)	654 (3 studies)	⊕⊖⊖⊖ **very low**[2,3,4]	
	Moderate					
	113 per 1000	**193 per 1000** (153 to 243)				
Response (CBASP + TAU/nefazodone versus TAU/nefazodone) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)/Inventory of Depressive Symptoms (IDS) Follow-up: 12-52 weeks	**Study population**		**RR 1.35** (1 to 1.83)	585 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,4]	
	195 per 1000	**264 per 1000** (195 to 357)				
	Moderate					
	204 per 1000	**275 per 1000** (204 to 373)				
Depression symptomatology (MBCT+TAU versus TAU) Hamilton Rating Scale for Depression (HAM-D; change score)/Beck Depression Inventory (BDI-II; change score) Follow-up: 8-12 weeks		The mean depression symptomatology (MBCT+TAU versus TAU) in the intervention groups was **1.14 standard deviations lower** (2.1 to 0.19 lower)		117 (3 studies)	⊕⊖⊖⊖ **very low**[5,6,7]	SMD -1.14 (-2.1 to -0.19)
Depression symptomatology (CBASP + TAU/nefazodone versus TAU/nefazodone) Hamilton Rating Scale for Depression (HAM-D; change score)/Inventory of Depressive Symptoms (IDS; change score) Follow-up: 8-52 weeks		The mean depression symptomatology (CBASP + TAU/nefazodone versus TAU/nefazodone) in the intervention groups was **0.8 standard deviations lower** (1.13 to 0.47 lower)		610 (3 studies)	⊕⊖⊖⊖ **very low**[8,9]	SMD -0.8 (-1.13 to -0.47)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	Cognitive and cognitive behavioural therapies + TAU/AD				
Depression symptomatology (CBT [group] + TAU versus waitlist + TAU) Beck Depression Inventory (BDI; change score) Follow-up: mean 10 weeks		The mean depression symptomatology (CBT [group] + TAU versus waitlist + TAU) in the intervention groups was **0.85 standard deviations lower** (1.29 to 0.41 lower)		88 (1 study)	⊕⊖⊖⊖ **very low**[7,10]	SMD -0.85 (-1.29 to -0.41)
Discontinuation for any reason (MBCT+TAU versus TAU) Number of participants discontinuing for any reason including adverse events Follow-up: 8-12 weeks	**Study population**		**RR 2.85** (0.84 to 9.66)	130 (3 studies)	⊕⊕⊖⊖ **low**[11,12]	
	62 per 1000	**178 per 1000** (52 to 604)				
	Moderate					
	67 per 1000	**191 per 1000** (56 to 647)				
Discontinuation for any reason (CBASP + TAU/nefazodone versus TAU/nefazodone) Number of participants discontinuing for any reason including adverse events Follow-up: 8-52 weeks	**Study population**		**RR 1.09** (0.54 to 2.17)	662 (3 studies)	⊕⊖⊖⊖ **very low**[9,13,14]	
	237 per 1000	**259 per 1000** (128 to 515)				
	Moderate					
	261 per 1000	**284 per 1000** (141 to 566)				
Discontinuation for any reason (CBT [group] + TAU versus waitlist + TAU) Number of participants discontinuing for any reason including adverse events Follow-up: mean 10 weeks	**Study population**		**RR 0.06** (0 to 0.99)	96 (1 study)	⊕⊕⊖⊖ **low**[2,15]	
	167 per 1000	**10 per 1000** (0 to 165)				
	Moderate					
	167 per 1000	**10 per 1000** (0 to 165)				
Discontinuation due to adverse events (CBASP + nefazodone versus nefazodone)	**Study population**		**RR 0.51** (0.29 to 0.91)	453 (1 study)	⊕⊖⊖⊖ **very low**[2,4,16]	
	137 per 1000	**70 per 1000** (40 to 125)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	Cognitive and cognitive behavioural therapies + TAU/AD				
Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Moderate** **137 per 1000**	**70 per 1000** (40 to 125)				

Notes:
*The basis for the assumed risk (e.g. 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).
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline. Non-blind intervention administrator(s) and participants, although outcome assessors are blind. Unclear risk of attrition bias (>20% difference in drop-out between groups but ITT analysis used)
[2] Events<300
[3] Non-blind intervention administrator(s) or participants, although outcome assessors are blinded. Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[4] Funding from pharmaceutical company
[5] Non-blind intervention administrator(s) and participants and outcome assessment either non-blind or blinding unclear
[6] I-squared=>80%
[7] N<400
[8] High risk of bias associated with randomisation method due to significant difference between groups at baseline in studies contributing >50% to analysis. Non-blind intervention administrator(s) and participants, although outcome assessors are blind. Unclear risk of attrition bias (drop-out>20% or difference between groups>20%)
[9] I-squared>50%
[10] Unclear randomisation method and method of allocation concealment. Non-blind intervention administration and outcome assessment
[11] Unclear (or high risk associated with) randomisation method, and non-blind intervention administrator(s) and participants
[12] 95% CI crosses both the line of no effect and the threshold for clinically important harm (RR 1.25)
[13] High risk of bias associated with randomisation method due to significant difference between groups at baseline in studies contributing >50% to analysis. Non-blind intervention administrator(s) and participants
[14] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[15] Unclear randomisation method and method of allocation concealment and non-blind intervention administrator(s) and participants
[16] Non-blind intervention administrator(s) and participants

Update 2017

Table 153: Study information table for trials included in the meta-analysis of cognitive and cognitive behavioural therapies versus attention-placebo for relapse prevention

	CBASP (maintenance treatment) versus assessment only
Total no. of studies (N randomised)	1 (82)
Study ID	Klein 2004
Country	US

	CBASP (maintenance treatment) versus assessment only
Chronic definition	Mixed (39% chronic major depression , 39% double depression and 22% recurrent depression with incomplete remission between episodes)
Age range (mean)	Range NR (45.1)
Sex (% female)	67
Ethnicity (% BME)	8
Mean age (SD) at first onset of depression	28.2 (12.9)
Mean months (SD) since onset of current episode	88.8 (117.6)
No. (SD) of previous depressive episodes	2.4 (1.6)
Previous treatment	65% psychotherapy; 60% antidepressants; 45% both antidepressants and psychotherapy; 20% no prior treatment for depression; AND acute phase or cross-over treatment with CBASP (Keller 2000)
Baseline severity	HAMD 6.4 (Less severe)
Intervention details	CBASP (maintenance treatment; followed the manual by McCullough 2000)
Intervention dose	13 sessions (1 every 4 weeks; mean attended 11.1 sessions [SD=3.8])
Comparator details (mean dose, if applicable)	Assessment-only (13 sessions [1 every 4 weeks])
Treatment length (weeks)	52
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation	

Update 2017

Table 154: Summary of findings table for cognitive and cognitive behavioural therapies compared to attention-placebo for relapse prevention

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Assessment-only	CBASP (maintenance treatment)				
Relapse Number of people scoring ≥16 on Hamilton Rating Scale for Depression (HAM-D) on 2 consecutive visits AND meeting DSM-IV criteria for a diagnosis of MDD Follow-up: mean 52 weeks	**Study population**		**RR 0.12** (0.02 to 0.91)	82 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	200 per 1000	**24 per 1000** (4 to 182)				
	Moderate					
	200 per 1000	**24 per 1000** (4 to 182)				

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Assessment-only	CBASP (maintenance treatment)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 52 weeks		The mean depression symptomatology in the intervention groups was **0.91 standard deviations lower** (1.37 to 0.45 lower)		82 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	SMD -0.91 (-1.37 to -0.45)
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 52 weeks	**Study population**		**RR 0.87** (0.41 to 1.81)	82 (1 study)	⊕⊖⊖⊖ **very low**[3,5,6]	
	275 per 1000	**239 per 1000** (113 to 498)				
	Moderate					
	275 per 1000	**239 per 1000** (113 to 498)				

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and method of allocation concealment is unclear. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded. Unclear risk of attrition bias (drop-put>20% but difference between groups <20% and ITT analysis used)
[2] Events<300
[3] Funding from pharmaceutical company
[4] N<400
[5] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and method of allocation concealment is unclear. Non-blind intervention administrator(s) and participants
[6] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Update 2017

Table 155: Study information table for trials included in the meta-analysis of IPT versus antidepressants

	IPT versus sertraline
Total no. of studies (N randomised)	1 (47)
Study ID	Markowitz 2005
Country	US
Chronic definition	DSM-IV early-onset (<21 years) dysthymic disorder (confirmed with SCID)
Age range (mean)	NR by arm (for all four arms of study: Range NR [42.3])
Sex (% female)	NR by arm (for all four arms of study: 63)
Ethnicity (% BME)	NR by arm (for all four arms of study: 37)
Mean age (SD) at first onset of depression	NR (inclusion criteria <21 years)

	IPT versus sertraline
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	HAMD 18.3 (Less severe)
Intervention details	IPT for dysthymic disorder (IPT-D; followed manual by Markowitz 1998)
Intervention dose	16-18 x 50-min sessions (mean attended 13.2 sessions [SD=4.0])
Comparator details (mean dose, if applicable)	Sertraline 50-200mg/day (mean daily dose 111.9 mg/day [SD=56.3]) + clinical management (10 x clinical management sessions [mean attended 7.5 sessions [SD=3.3])
Treatment length (weeks)	16
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation Markowitz 2005 is a four-armed trial but, where possible, data is extracted for only the two relevant arms here	

1 **Table 156: Summary of findings table for IPT compared to antidepressants**

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Sertraline	IPT				
Remission Number of people scoring <7 on Hamilton Rating Scale for Depression (HAM-D) AND >50% improvement on HAMD AND GAF score>70 Follow-up: mean 16 weeks	**Study population**		**RR 0.52** (0.21 to 1.29)	47 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	417 per 1000	**217 per 1000** (87 to 538)				
	Moderate					
	417 per 1000	**217 per 1000** (88 to 538)				
Response Number of people showing ≥40% improvement on Montgomery Asberg Depression Rating Scale (MADRS)/≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 16-26 weeks	**Study population**		**RR 0.76** (0.63 to 0.92)	421 (2 studies)	⊕⊖⊖⊖ **very low**[3,4,5]	
	595 per 1000	**453 per 1000** (375 to 548)				
	Moderate					
	590 per 1000	**448 per 1000** (372 to 543)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery		The mean depression symptomatology in the intervention groups was **0.49 standard**		421 (2 studies)	⊕⊖⊖⊖ **very low**[4,6]	SMD 0.49 (0.24 to 0.74)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Sertraline	IPT				
Asberg Depression Rating Scale (MADRS; change score) Follow-up: 16-26 weeks		**deviations higher** (0.24 to 0.74 higher)				
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 16 weeks	**Study population**		**RR 0.83** (0.26 to 2.73)	47 (1 study)	⊕⊖⊖⊖ **very low**[1,3,7]	
	208 per 1000	**173 per 1000** (54 to 569)				
	Moderate					
	208 per 1000	**173 per 1000** (54 to 568)				

[1] Unclear randomisation and method of allocation concealment. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (Rr 0.75) and clinically important benefit (RR 1.25)
[3] Study partially funded by pharmaceutical company
[4] High risk of bias associated with randomisation bias due to significant difference between groups at baseline. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded
[5] Events<300
[6] N<400
[7] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Update 2017

Table 157: Study information table for trials included in the meta-analysis of IPT versus other psychological interventions

	IPT versus brief supportive psychotherapy (BSP)
Total no. of studies (N randomised)	1 (49)
Study ID	Markowitz 2005
Country	US
Chronic definition	DSM-IV early-onset (<21 years) dysthymic disorder (confirmed with SCID)
Age range (mean)	NR by arm (for all four arms of study: Range NR [42.3])
Sex (% female)	NR by arm (for all four arms of study: 63)
Ethnicity (% BME)	NR by arm (for all four arms of study: 37)
Mean age (SD) at first onset of depression	NR (inclusion criteria <21 years)
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	HAMD 19.3 (Less severe)
Intervention details	IPT for dysthymic disorder (IPT-D; followed manual by Markowitz 1998)

	IPT versus brief supportive psychotherapy (BSP)
Intervention dose	16-18 x 50-min sessions (mean attended 13.2 sessions [SD=4.0])
Comparator details (mean dose, if applicable)	Brief supportive psychotherapy (BSP). 16-18 x 50-min sessions (mean attended 9.6 sessions [SD=6.3])
Treatment length (weeks)	16
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation Markowitz 2005 is a four-armed trial but, where possible, data is extracted for only the two relevant arms here	

Table 158: Summary of findings table for IPT compared to other psychological 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	IPT versus brief supportive psychotherapy (BSP) for dysthymia				
Remission Number of people scoring <7 on Hamilton Rating Scale for Depression (HAM-D) AND >50% improvement on HAMD AND GAF score>70 Follow-up: mean 16 weeks	**Study population**		**RR 1.88** (0.5 to 7.03)	49 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	115 per 1000	**217 per 1000** (58 to 811)				
	Moderate					
	115 per 1000	**216 per 1000** (58 to 808)				
Response Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 16 weeks	**Study population**		**RR 1.13** (0.51 to 2.52)	49 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	308 per 1000	**348 per 1000** (157 to 775)				
	Moderate					
	308 per 1000	**348 per 1000** (157 to 776)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 16 weeks		The mean depression symptomatology in the intervention groups was **0.06 standard deviations lower** (0.63 lower to 0.5 higher)		49 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	SMD -0.06 (-0.63 to 0.5)

Update 2017

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	IPT versus brief supportive psychotherapy (BSP) for dysthymia				
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 16 weeks	**Study population**		**RR 0.41** (0.15 to 1.11)	49 (1 study)	⊕⊖⊖⊖ **very low**[3,5,6]	
	423 per 1000	**173 per 1000** (63 to 470)				
	Moderate					
	423 per 1000	**173 per 1000** (63 to 470)				

Notes:
[1] Randomisation method and method of allocation concealment unclear. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded. Unclear risk of attrition bias (drop-out>20% and difference between groups>20% but ITT analysis used)
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] Study partially funded by pharmaceutical company
[4] 95% CI crosses line of no effect and threshold for both clinically important benefit (SMD -0.5) and clinically important harm (SMD 0.5)
[5] Randomisation method and method of allocation concealment unclear. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded.
[6] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 0.75)

Update 2017

Table 159: Study information table for trials included in the meta-analysis of IPT in combination with antidepressants or treatment as usual versus antidepressants or treatment as usual-only

	IPT + TAU/AD versus TAU/AD-only
Total no. of studies (N randomised)	4 (654)
Study ID	Browne 2002[1] de Mello 2001[2] Markowitz 2005[3] Schramm 2008[4]
Country	Canada[1] Brazil[2] US[3] Germany[4]
Chronic definition	Dysthymia[1] Double depression (91%; + 9% dysthymic disorder) [2] Dysthymia (early-onset [<21 years]) [3] Double depression (53%; + 47% chronic MDD ≥2 years) [4]
Age range (mean)	Range NR (41.9) [1] NR[2] NR by arm (for all four arms of study: Range NR [42.3]) [3] Range NR (42.8) [4]

	IPT + TAU/AD versus TAU/AD-only
Sex (% female)	NR by arm (for all three arms of study: 68) [1] 80[2] NR by arm (for all four arms of study: 63) [3] 67[4]
Ethnicity (% BME)	NR[1,2,4] NR by arm (for all four arms of study: 37) [3]
Mean age (SD) at first onset of depression	NR[1,2] NR (inclusion criteria <21 years) [3] NR (27% early onset) [4]
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR[1,2,3] Mean NR (29% 1 episode; 71% ≥2 episodes) [4]
Previous treatment	NR[1,2,3] 76% psychotherapy; 76% pharmacotherapy; 51% hospitalization[4]
Baseline severity	MADRS 25.5 (Less severe) [1] MADRS 19.4 (Less severe) [2] HAMD 18.7 (Less severe) [3] HAMD 24.6 (More severe) [4]
Intervention details	IPT (followed the manual by Weissman and Klerman 1993 and Klerman et al. 1984) + sertraline[1] IPT (adapted to dysthymic disorder) + moclobemide[2] IPT for dysthymic disorder (IPT-D; followed manual by Markowitz 1998) + sertraline[3] IPT (followed the modified version of the original IPT manual by Klerman et al. 1984 for use in an inpatient setting, Schramm 2001) + standard pharmacotherapy (sertraline or, as the second line treatment, amitriptyline or amitriptyline-N-oxide) [4]
Intervention dose	12x 1-hour sessions (mean attended 8.9 sessions [SD=2.6]) + 50-200,g/day of sertraline[1] 16 sessions + 300-600mg/day (mean dose 460.71 mg/day [SD=124.71]) [2] 16-18 x 50-min sessions (mean attended 12.8 sessions [SD=4.01]) + 50-200mg/day (mean daily dose 116.3 mg/day [SD=43.9]) [3] 15x individual sessions (mean attended 11.54 sessions [SD=3.43] and 8 additional IPT-group sessions + sertraline (mean final dose 80.2 mg/day [SD=32.9]), amitriptyline or amitriptyline-N-oxide (mean final dose 160.8 mg/day [SD=58.2]) [4]
Comparator details (mean dose, if applicable)	Sertraline 50-200mg/day[1] Moclobemide 300-600mg/day (mean dose 490.90 mg/day [SD=117.93]) + clinical management[2] Sertraline 50-200mg/day (mean daily dose 111.9 mg/day [SD=56.3]) + clinical management (10 x clinical management sessions [mean attended 7.5 sessions, SD=3.3]) [3] Sertraline (mean final dose 80.2 mg/day [SD=32.9]), amitriptyline or amitriptyline-N-oxide (mean final dose 160.8 mg/day [SD=58.2]) + 15x 15-20-min sessions of clinical management[4]

Update 2017

	IPT + TAU/AD versus TAU/AD-only
Treatment length (weeks)	26[1] 12[2] 16[3] 5[4]

Notes:
Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation
[1]Browne 2002; [2]de Mello 2001; [3]Markowitz 2005; [4]Schramm 2008
Browne 2002 is a three-armed trial and Markowitz 2005 is a four-armed trial but, where possible, data is extracted for only the two relevant arms here

Table 160: Summary of findings table for IPT in combination with antidepressants or treatment as usual compared to antidepressants or treatment as usual-only

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	IPT + TAU/AD				
Remission (IPT + TAU/AD versus TAU/AD-only) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D)/<7 on HAMD-D >50% improvement on HAMD AND GAF score>70 Follow-up: 5-16 weeks	**Study population** **356 per 1000**	**508 per 1000** (313 to 828) **Moderate** **502 per 1000** (309 to 818)	**RR 1.43** (0.88 to 2.33)	90 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
Remission (IPT + standard pharmacotherapy versus standard pharmacotherapy + clinical management) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 5 weeks	**Study population** **286 per 1000**	**500 per 1000** (229 to 1000) **Moderate** **500 per 1000** (229 to 1000)	**RR 1.75** (0.8 to 3.84)	45 (1 study)	⊕⊕⊖⊖ **low**[2,4]	
Remission (IPT + sertraline versus sertraline) Number of people scoring <7 on Hamilton Rating Scale for Depression (HAM-D) AND >50% improvement on HAMD AND GAF score>70 Follow-up: mean 16 weeks	**Study population** **417 per 1000**	**525 per 1000** (279 to 979) **Moderate** **525 per 1000** (279 to 980)	**RR 1.26** (0.67 to 2.35)	45 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	IPT + TAU/AD				
Response (IPT + TAU/AD versus TAU/AD-only) Number of people showing ≥40% improvement on Montgomery Asberg Depression Rating Scale (MADRS)/≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 5-26 weeks	**Study population**		**RR 1.11** (0.79 to 1.56)	498 (3 studies)	⊕⊖⊖⊖ **very low**[2,3,6,7]	
	577 per 1000	**640 per 1000** (456 to 900)				
	Moderate					
	583 per 1000	**647 per 1000** (461 to 909)				
Response (IPT + standard pharmacotherapy versus standard pharmacotherapy + clinical management) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 5 weeks	**Study population**		**RR 1.86** (1.02 to 3.4)	45 (1 study)	⊕⊕⊖⊖ **low**[4,8]	
	381 per 1000	**709 per 1000** (389 to 1000)				
	Moderate					
	381 per 1000	**709 per 1000** (389 to 1000)				
Response (IPT + sertraline versus sertraline) Number of people showing ≥40% improvement on Montgomery Asberg Depression Rating Scale (MADRS)/≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 16-26 weeks	**Study population**		**RR 0.97** (0.83 to 1.13)	453 (2 studies)	⊕⊖⊖⊖ **very low**[6,8,9]	
	595 per 1000	**578 per 1000** (494 to 673)				
	Moderate					
	590 per 1000	**572 per 1000** (490 to 667)				
Depression symptomatology (IPT + TAU/AD versus TAU/AD-only) Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery Asberg Depression Rating Scale (MADRS;		The mean depression symptomatology (IPT + TAU/AD versus TAU/AD-only) in the intervention groups was **0.16 standard deviations lower** (0.43 lower to 0.11 higher)		522 (4 studies)	⊕⊖⊖⊖ **very low**[6,9]	SMD -0.16 (-0.43 to 0.11)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	IPT + TAU/AD				
change score) Follow-up: 5-26 weeks						
Depression symptomatology (IPT + standard pharmacotherapy versus standard pharmacotherapy + clinical management) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 5 weeks		The mean depression symptomatology (IPT + standard pharmacotherapy versus standard pharmacotherapy + clinical management) in the intervention groups was **0.71 standard deviations lower** (1.32 to 0.1 lower)		45 (1 study)	⊕⊕⊖⊖ **low**[4,10]	SMD -0.71 (-1.32 to -0.1)
Depression symptomatology (IPT + moclobemide versus moclobemide + clinical management) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: mean 12 weeks		The mean depression symptomatology (IPT + moclobemide versus moclobemide + clinical management) in the intervention groups was **0.03 standard deviations lower** (0.83 lower to 0.77 higher)		24 (1 study)	⊕⊖⊖⊖ **very low**[11,12]	SMD -0.03 (-0.83 to 0.77)
Depression symptomatology (IPT + sertraline versus sertraline) Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 16-26 weeks		The mean depression symptomatology (IPT + sertraline versus sertraline) in the intervention groups was **0.06 standard deviations lower** (0.24 lower to 0.12 higher)		453 (2 studies)	⊕⊖⊖⊖ **very low**[6,9]	SMD -0.06 (-0.24 to 0.12)
Discontinuation for any reason (IPT + TAU/AD versus TAU/AD-only) Number of participants discontinuing for any reason including adverse events Follow-up: 5-16 weeks	**Study population** **281 per 1000** **Moderate** **208 per 1000**	**267 per 1000** (127 to 560) **198 per 1000** (94 to 414)	**RR 0.95** (0.45 to 1.99)	125 (3 studies)	⊕⊖⊖⊖ **very low**[13,14]	
	Study population					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	IPT + TAU/AD				
Discontinuation for any reason (IPT + standard pharmacotherapy versus standard pharmacotherapy + clinical management) Number of participants discontinuing for any reason including adverse events Follow-up: mean 5 weeks	**95 per 1000**	**250 per 1000** (56 to 1000)	**RR 2.62** (0.59 to 11.64)	45 (1 study)	⊕⊖⊖⊖ **very low**[4,14]	
	Moderate					
	95 per 1000	**249 per 1000** (56 to 1000)				
Discontinuation for any reason (IPT + moclobemide versus moclobemide + clinical management) Number of participants discontinuing for any reason including adverse events Follow-up: mean 12 weeks	**Study population**		**RR 0.65** (0.31 to 1.36)	35 (1 study)	⊕⊖⊖⊖ **very low**[13,14]	
	579 per 1000	**376 per 1000** (179 to 787)				
	Moderate					
	579 per 1000	**376 per 1000** (179 to 787)				
Discontinuation for any reason (IPT + sertraline versus sertraline) Number of participants discontinuing for any reason including adverse events Follow-up: mean 16 weeks	**Study population**		**RR 0.91** (0.28 to 2.97)	45 (1 study)	⊕⊖⊖⊖ **very low**[3,6,14]	
	208 per 1000	**190 per 1000** (58 to 619)				
	Moderate					
	208 per 1000	**189 per 1000** (58 to 618)				

Notes:
[1] Randomisation method is unclear and unclear method of allocation concealment. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Study partially funded by pharmaceutical company
[4] Baseline group comparability is unclear and unclear method of allocation concealment. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded
[5] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[6] High risk associated with randomisation method due to significant difference between groups at baseline. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded
[7] I-squared>50%
[8] Events<300
[9] Data cannot be extracted or is not reported for all outcomes and study partially funded by

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	TAU/AD-only	IPT + TAU/AD				

pharmaceutical company
[10] N<400
[11] Unclear randomisation method and method of allocation concealment. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded. High risk of attrition bias (drop-out>20% and difference between groups>20%)
[12] 95% CI crosses line of no effect and threshold for both clinically important benefit (SMD -0.5) and clinically important harm (SMD 0.5)
[13] Unclear randomisation method and method of allocation concealment. Non-blind intervention administrator(s) and participants
[14] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 161: Study information table for trials included in the meta-analysis of brief supportive psychotherapy versus antidepressants

	Brief supportive psychotherapy (BSP) versus sertraline
Total no. of studies (N randomised)	1 (50)
Study ID	Markowitz 2005
Country	US
Chronic definition	DSM-IV early-onset (<21 years) dysthymic disorder (confirmed with SCID)
Age range (mean)	NR by arm (for all four arms of study: Range NR [42.3])
Sex (% female)	NR by arm (for all four arms of study: 63)
Ethnicity (% BME)	NR by arm (for all four arms of study: 37)
Mean age (SD) at first onset of depression	NR (inclusion criteria <21 years)
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	HAMD 18.8 (Less severe)
Intervention details	Brief supportive psychotherapy (BSP)
Intervention dose	16-18 x 50-min sessions (mean attended 9.6 sessions [SD=6.3])
Comparator details (mean dose, if applicable)	Sertraline 50-200mg/day (mean daily dose 111.9 mg/day [SD=56.3]) + clinical management (10 x clinical management sessions [mean attended 7.5 sessions [SD=3.3])
Treatment length (weeks)	16
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation Markowitz 2005 is a four-armed trial but, where possible, data is extracted for only the two relevant arms here	

Update 2017

Table 162: Summary of findings table for brief supportive psychotherapy compared to antidepressants

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Sertraline	Corresponding risk Brief supportive psychotherapy (BSP)				
Remission Number of people scoring <7 on Hamilton Rating Scale for Depression (HAM-D) AND >50% improvement on HAMD AND GAF score>70 Follow-up: mean 16 weeks	**Study population**		**RR 0.28** (0.09 to 0.89)	50 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	417 per 1000	**117 per 1000** (38 to 371)				
	Moderate					
	417 per 1000	**117 per 1000** (38 to 371)				
Response Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 16 weeks	**Study population**		**RR 0.53** (0.27 to 1.03)	50 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	583 per 1000	**309 per 1000** (158 to 601)				
	Moderate					
	583 per 1000	**309 per 1000** (157 to 600)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 16 weeks		The mean depression symptomatology in the intervention groups was **0.77 standard deviations higher** (0.19 to 1.34 higher)		50 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	SMD 0.77 (0.19 to 1.34)
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 16 weeks	**Study population**		**RR 2.03** (0.83 to 4.99)	50 (1 study)	⊕⊖⊖⊖ **very low**[3,6,7]	
	208 per 1000	**423 per 1000** (173 to 1000)				
	Moderate					
	208 per 1000	**422 per 1000** (173 to 1000)				

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and method of allocation concealment is unclear. Non-blind intervention administrator(s) and participants, although outcome assessors are blinded. High risk of attrition bias (>20% drop-out and difference between groups >20%), although ITT analysis used
[2] Events<300
[3] Study partially funded by pharmaceutical company

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Sertraline	Brief supportive psychotherapy (BSP)				

[4] No explanation was provided
[5] N<400
[6] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and method of allocation concealment is unclear. Non-blind intervention administrator(s) and participants
[7] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)

Table 163: Study information table for trials included in the meta-analysis of Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD) combined with antidepressants versus antidepressants-only for relapse prevention

	Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD) + fluoxetine versus fluoxetine (maintenance treatment)
Total no. of studies (N randomised)	1 (40)
Study ID	Hellerstein 2001
Country	US
Chronic definition	DSM-III-R early-onset (<21 years) dysthymia
Age range (mean)	Range NR (45.10)
Sex (% female)	50
Ethnicity (% BME)	13
Mean age (SD) at first onset of depression	NR (inclusion criteria <21 years)
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	3 (2.51)
Previous treatment	80% had had previous individual psychotherapy (average number of months in therapy was 27.75 [SD=25.99]) and 25% previous group therapy experience; AND 8-week acute treatment phase with fluoxetine (10-80mg/day; partial responders randomized for relapse prevention phase)
Baseline severity	HAMD 7 (Less severe)
Intervention details	Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD; followed an unpublished manual) + fluoxetine
Intervention dose	16x weekly 1.5-hour sessions + 20-80mg/day of fluoxetine (mean final dose 37.36mg [SD=17.27])
Comparator details (mean dose, if applicable)	Fluoxetine (maintenance treatment). 20-80mg/day (mean final dose 38.75mg [SD=18.93])
Treatment length (weeks)	16
Notes:	

Update 2017

	Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD) + fluoxetine versus fluoxetine (maintenance treatment)

Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation

Table 164: Summary of findings table for Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD) combined with antidepressants compared to antidepressants-only for relapse prevention

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 Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD) + fluoxetine versus fluoxetine (maintenance treatment) for relapse prevention in dysthymia				
Relapse Number of people scoring >0 on item #1 (depressed mood) on Hamilton Rating Scale for Depression (HAM-D) OR meeting DSM-IV criteria for a diagnosis of dysthymia Follow-up: mean 16 weeks	**Study population** **375 per 1000**	**176 per 1000** (53 to 589)	**RR 0.47** (0.14 to 1.57)	33 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	Moderate **375 per 1000**	**176 per 1000** (53 to 589)				
Response Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND much/very much improved on CGI-I (score 1-2) Follow-up: mean 16 weeks	**Study population** **765 per 1000**	**887 per 1000** (650 to 1000)	**RR 1.16** (0.85 to 1.59)	35 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	
	Moderate **765 per 1000**	**887 per 1000** (650 to 1000)				
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 16 weeks	**Study population** **150 per 1000**	**101 per 1000** (18 to 535)	**RR 0.67** (0.12 to 3.57)	40 (1 study)	⊕⊖⊖⊖ **very low**[2,3,5]	
	Moderate **150 per 1000**	**101 per 1000** (18 to 535)				

Update 2017

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	**Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD) + fluoxetine versus fluoxetine (maintenance treatment) for relapse prevention in dysthymia**				

Notes:
[1] Unclear randomisation method and method of allocation concealment. Non-blind intervention administrator(s) and participants and blinding of outcome assessors is unclear
[2] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)
[3] Funding from pharmaceutical company
[4] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[5] Unclear randomisation method and method of allocation concealment. Non-blind intervention administrator(s) and participants

Update 2017

9.3.21 Pharmacological interventions for chronic depression

Evidence was found for 11 comparisons of pharmacological interventions as follows: SSRIs compared to placebo (see Table 165 for study characteristics); SSRIs compared to other pharmacological interventions (see Table 167 for study characteristics); SSRIs combined with a psychological intervention compared to a psychological intervention only (see Table 169 for study characteristics); TCAs compared to placebo (see Table 171 for study characteristics); TCAs compared to other pharmacological interventions (see Table 173 for study characteristics); SNRIs compared to placebo (see Table 175 for study characteristics); MAOIs compared to placebo (see Table 177 for study characteristics); MAOIs compared to other pharmacological interventions (see Table 179 for study characteristics); reversible inhibitors of monoamine oxidase (RIMAs) compared to placebo (see Table 181 for study characteristics); RIMAs compared to other pharmacological interventions (see Table 183 for study characteristics); antipsychotics compared to placebo (see Table 185 for study characteristics).

Evidence for these comparisons are summarised in the clinical GRADE evidence profiles below (Table 166, Table 168, Table 170, Table 172, Table 174, Table 176, Table 178, Table 180, Table 182, Table 184 and Table 186). See also the full GRADE evidence profiles in Appendix L, forest plots in Appendix M and the full study characteristics, comparisons and outcomes tables in Appendix J6.

Table 165: Study information table for trials included in the meta-analysis of SSRIs versus placebo

	SSRIs versus placebo
Total no. of studies (N randomised)	7 (671)
Study ID	Anisman 1999[1] Hellerstein 1993[2] Hellerstein 2010[3]

	SSRIs versus placebo
	Ravindran 2013[4] Thase 1996[5] Vanelle 1997[6] Williams 2000[7]
Country	Canada[1,4] US[2,3,5,7] France[6]
Chronic definition	Dysthymia[1,3,4,6] Early-onset (<21 years) dysthymia[2,5] Dysthymia (trial also included minor depression but data only extracted for subgroup with dysthymia)[7]
Age range (mean)	Range NR (40.5)[1] Range NR (36.2)[2] 23-65 (44.7)[3] 19-59 (41.5)[4] Range NR (42.1)[5] Range NR (43)[6] NR[7]
Sex (% female)	51[1] 50[2,3] 48[4] 64[5] 75[6] NR[7]
Ethnicity (% BME)	NR[1,2,6,7] 28[3] 8[4] 5[5]
Mean age (SD) at first onset of depression	NR[1,7] NR (inclusion criteria <21 years: by self-report 62.5% began in childhood, 25% in teens and 12.5% in early 20s)[2] NR (75% had early-onset dysthymic disorder)[3] 25.8 (12.9)[4] 12.2 (4.8)[5] NR (23% early-onset and 77% late-onset dysthymia)[6]
Mean months (SD) since onset of current episode	NR[1,2,3,7] 223.8 (140.2)[4] 359.8 (127.9)[5] 73.0 (SD NR)[6]
No. (SD) of previous depressive episodes	NR[1,2,4,5,6,7] Mean NR (39% no previous major depressive episodes, 19% 1 prior major depression, and 42% ≥2 earlier episodes of major depression)[3]
Previous treatment	NR[1,3,4,5,7] 88% previous psychotherapy; 19% current psychotherapy; 13% prior antidepressant response[2] 17% current psychotherapy; 48% previous psychotropic treatment, 59% current benzodiazepine use[6]
Baseline severity	HAMD 17.8 (Less severe)[1]

Update 2017

	SSRIs versus placebo
	HAMD 19 (Less severe)[2] HAMD 23.4 (Less severe)[3] HAMD 18.8 (Less severe)[4] HAMD 12.7 (Less severe)[5] HAMD 20.6 (Less severe)[6] NR[7]
Intervention details	Sertraline[1,5] Fluoxetine[2,6] Escitalopram[3] Paroxetine[4] (+ clinical management)[7]
Intervention dose	50-200mg/day[1] 20mg/day (actual doses taken 10-60mg/day; mean final dose 32.7mg [SD=13.8])[2] 10-20mg/day (mean final dose 15.3mg [SD=5.1])[3] 10-40mg/day (mean final dose 33.33 mg/day)[4] 50-200mg/day (mean final dose 139.6mg [SD=58.5])[5] 20mg/day[6] 10-40mg/day (+ 6x 15-min sessions of medication management)[7]
Comparator details (mean dose, if applicable)	Placebo[1,2,5,6] Placebo 10-20mg/day (mean final dose 16.7 mg [SD=4.9])[3] Placebo 10-40mg/day (mean final dose 35.25 mg/day)[4] Placebo 10-40mg/day (+ 6x 15-min sessions of medication management)[7]
Treatment length (weeks)	12[1,3,4,5] 8[2] 13[6] 11[7]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Anisman 1999; [2]Hellerstein 1993; [3]Hellerstein 2010; [4]Ravindran 2013; [5]Thase 1996; [6]Vanelle 1997; [7]Williams 2000 Thase 1996[5] and Williams 2000[7] are three-armed trials but, where possible, data is extracted for only the two relevant arms here. From Williams 2000[7] data also only extracted for dysthymia subgroup and as a result demographic details limited (not reported by diagnostic subgroup)	

Update 2017

1 **Table 166: Summary of findings table for SSRIs compared to 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				
	Placebo	SSRIs				
Remission (any SSRI) Number of people scoring <7/≤4/7/8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 11-13 weeks	**Study population**		**RR 1.47** (1.15 to 1.87)	578 (5 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	307 per 1000	**451 per 1000** (353 to 574)				
	Moderate					

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	SSRIs				
	256 per 1000	**376 per 1000** (294 to 479)				
Remission (sertraline) Number of people scoring ≤4 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 12 weeks	**Study population**		**RR 1.46** (1.08 to 1.98)	274 (1 study)	⊕⊖⊖⊖ **very low**[2,3,4]	
	321 per 1000	**469 per 1000** (347 to 636)				
	Moderate					
	321 per 1000	**469 per 1000** (347 to 636)				
Remission (fluoxetine) Number of people scoring ≤7 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 13 weeks	**Study population**		**RR 1.73** (0.96 to 3.14)	111 (1 study)	⊕⊖⊖⊖ **very low**[5,6,7]	
	256 per 1000	**444 per 1000** (246 to 805)				
	Moderate					
	256 per 1000	**443 per 1000** (246 to 804)				
Remission (escitalopram) Number of people scoring ≤4 on Hamilton Rating Scale for Depression (HAM-D) AND HAMD item # 1 (depressed mood) score=0 Follow-up: mean 12 weeks	**Study population**		**RR 4** (0.5 to 32.2)	34 (1 study)	⊕⊖⊖⊖ **very low**[3,5,6]	
	59 per 1000	**235 per 1000** (29 to 1000)				
	Moderate					
	59 per 1000	**236 per 1000** (30 to 1000)				
Remission (paroxetine) Number of people scoring <7/≤8 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 11-12 weeks	**Study population**		**RR 1.58** (0.68 to 3.66)	159 (2 studies)	⊕⊖⊖⊖ **very low**[8,9,10,11]	
	358 per 1000	**566 per 1000** (243 to 1000)				
	Moderate					
	307 per 1000	**485 per 1000** (209 to 1000)				
Response (any SSRI) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)	**Study population**		**RR 1.62** (1.29 to 2.03)	558 (6 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	379 per 1000	**614 per 1000** (489 to 769)				
	Moderate					

Update 2017

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	SSRIs				
AND HAMD score≤10/AND much/very much improved on CGI-I (score 1-2)/ AND/OR much/very much improved on CGI-I (score 1-2) Follow-up: 8-13 weeks	**309 per 1000**	**501 per 1000** (399 to 627)				
Response (sertraline) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND HAMD score≤10/Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 12 weeks	**Study population**		**RR 1.61** (0.99 to 2.64)	341 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,6,9]	
	416 per 1000	**670 per 1000** (412 to 1000)				
	Moderate					
	373 per 1000	**601 per 1000** (369 to 985)				
Response (fluoxetine) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND much/very much improved on CGI-I (score 1-2) Follow-up: 8-13 weeks	**Study population**		**RR 1.96** (1.05 to 3.64)	143 (2 studies)	⊕⊖⊖⊖ **very low**[2,5,7]	
	309 per 1000	**606 per 1000** (325 to 1000)				
	Moderate					
	273 per 1000	**535 per 1000** (287 to 994)				
Response (escitalopram) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND much/very much improved on CGI-I (score 1-2) Follow-up: mean 12 weeks	**Study population**		**RR 1.4** (0.55 to 3.55)	34 (1 study)	⊕⊖⊖⊖ **very low**[3,5,10]	
	294 per 1000	**412 per 1000** (162 to 1000)				
	Moderate					
	294 per 1000	**412 per 1000** (162 to 1000)				
Response (paroxetine) Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D)	**Study population**		**RR 2.11** (1.02 to 4.37)	40 (1 study)	⊕⊕⊖⊖ **low**[2,5]	
	316 per 1000	**666 per 1000** (322 to 1000)				
	Moderate					

Update 2017

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	SSRIs				
AND/OR much/very much improved on CGI-I (score 1-2) Follow-up: mean 12 weeks	**316 per 1000**	**667 per 1000** (322 to 1000)				
Depression symptomatology (any SSRI) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: 8-13 weeks		The mean depression symptomatology (any SSRI) in the intervention groups was **0.69 standard deviations lower** (1.02 to 0.35 lower)		556 (6 studies)	⊕⊖⊖⊖ **very low**[1,3,9]	SMD -0.69 (-1.02 to -0.35)
Depression symptomatology (sertraline) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 12 weeks		The mean depression symptomatology (sertraline) in the intervention groups was **0.61 standard deviations lower** (1.3 lower to 0.07 higher)		339 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,12,13]	SMD -0.61 (-1.3 to 0.07)
Depression symptomatology (fluoxetine) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: 8-13 weeks		The mean depression symptomatology (fluoxetine) in the intervention groups was **0.8 standard deviations lower** (1.81 lower to 0.21 higher)		143 (2 studies)	⊕⊖⊖⊖ **very low**[5,7,12,13]	SMD -0.8 (-1.81 to 0.21)
Depression symptomatology (escitalopram) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 12 weeks		The mean depression symptomatology (escitalopram) in the intervention groups was **0.9 standard deviations lower** (1.61 to 0.19 lower)		34 (1 study)	⊕⊖⊖⊖ **very low**[3,5,14]	SMD -0.9 (-1.61 to -0.19)
Depression symptomatology (paroxetine) Hamilton Rating Scale for		The mean depression symptomatology (paroxetine) in the		40 (1 study)	⊕⊕⊖⊖ **low**[5,14]	SMD -0.77 (-1.41 to -0.12)

Update 2017

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	SSRIs				
Depression (HAM-D; change score) Follow-up: mean 12 weeks		intervention groups was **0.77 standard deviations lower** (1.41 to 0.12 lower)				
Discontinuation for any reason (any SSRI) Number of participants discontinuing for any reason including adverse events Follow-up: 8-13 weeks	**Study population**		**RR 0.64** (0.42 to 0.96)	593 (6 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	215 per 1000	**137 per 1000** (90 to 206)				
	Moderate					
	223 per 1000	**143 per 1000** (94 to 214)				
Discontinuation for any reason (sertraline) Number of participants discontinuing for any reason including adverse events Follow-up: mean 12 weeks	**Study population**		**RR 0.62** (0.4 to 0.97)	342 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	241 per 1000	**150 per 1000** (97 to 234)				
	Moderate					
	239 per 1000	**148 per 1000** (96 to 232)				
Discontinuation for any reason (fluoxetine) Number of participants discontinuing for any reason including adverse events Follow-up: 8-13 weeks	**Study population**		**RR 1.17** (0.11 to 12.85)	175 (2 studies)	⊕⊖⊖⊖ **very low**[5,7,9,15]	
	200 per 1000	**234 per 1000** (22 to 1000)				
	Moderate					
	133 per 1000	**156 per 1000** (15 to 1000)				
Discontinuation for any reason (escitalopram) Number of participants discontinuing for any reason including adverse events Follow-up: mean 12 weeks	**Study population**		**RR 6.3** (0.35 to 113.81)	36 (1 study)	⊕⊖⊖⊖ **very low**[3,5,15]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				
Discontinuation for any reason (paroxetine) Number of participants discontinuing for any	**Study population**		**RR 0.68** (0.17 to 2.65)	40 (1 study)	⊕⊖⊖⊖ **very low**[5,15]	
	211 per 1000	**143 per 1000** (36 to 558)				

Update 2017

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	SSRIs				
reason including adverse events Follow-up: mean 12 weeks	**Moderate**					
	211 per 1000	**143 per 1000** (36 to 559)				
Discontinuation due to adverse events (any SSRI) Number of participants discontinuing due to adverse events Follow-up: 8-12 weeks	**Study population**		**RR 1.83** (0.69 to 4.86)	385 (4 studies)	⊕⊖⊖⊖ **very low**[1,3,15]	
	26 per 1000	**48 per 1000** (18 to 127)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				
Discontinuation due to adverse events (sertraline) Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Study population**		**RR 1.67** (0.56 to 4.98)	274 (1 study)	⊕⊖⊖⊖ **very low**[1,3,15]	
	36 per 1000	**60 per 1000** (20 to 178)				
	Moderate					
	36 per 1000	**60 per 1000** (20 to 179)				
Discontinuation due to adverse events (fluoxetine) Number of participants discontinuing due to adverse events Follow-up: mean 8 weeks	**Study population**		**RR 2.55** (0.11 to 58.6)	35 (1 study)	⊕⊖⊖⊖ **very low**[3,5,15]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				
Discontinuation due to adverse events (escitalopram) Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Study population**		**RR 2.7** (0.12 to 62.17)	36 (1 study)	⊕⊖⊖⊖ **very low**[3,5,15]	
	0 per 1000	**0 per 1000** (0 to 0)				
	Moderate					
	0 per 1000	**0 per 1000** (0 to 0)				
Discontinuation due to adverse events (paroxetine) Number of participants discontinuing due to adverse events	See comment	See comment	Not estimable	40 (1 study)	⊕⊕⊖⊖ **low**[2,5]	

Update 2017

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	SSRIs				
Follow-up: mean 12 weeks						

Notes:
[1] Unclear (or high risk of bias associated with) randomisation method and unclear method of allocation concealment. Unclear blinding of intervention administration and outcome assessment
[2] Events<300
[3] Funding from pharmaceutical company
[4] High risk of bias associated with randomisation method due to significant differences between groups at baseline and unclear method of allocation concealment. Unclear blinding of intervention administration and outcome assessment. Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[5] Unclear randomisation method and method of allocation concealment. Unclear blinding of intervention administration and outcome assessment
[6] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[7] Data is not reported for all outcomes
[8] Unclear blinding of intervention administrator(s)
[9] I-squared>50%
[10] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[11] Data is not reported for all outcomes and funding from pharmaceutical company
[12] I-squared>80%
[13] 95% CI crosses both line of no effect and threshold for clinically important benefit (SMD -0.5)
[14] N<400
[15] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 167: Study information table for trials included in the meta-analysis of SSRIs versus other pharmacological interventions

	Sertraline versus imipramine
Total no. of studies (N randomised)	2 (905)
Study ID	Keller 1998a[1] Thase 1996[2]
Country	US
Chronic definition	Double depression (54%; + 46% chronic MDD ≥2 years) [1] Early-onset (<21 years) dysthymia[2]
Age range (mean)	Range NR (41.1) [1] Range NR (41.8) [2]
Sex (% female)	63[1] 67[2]
Ethnicity (% BME)	9[1] 4[2]
Mean age (SD) at first onset of depression	MDD: 24.8 (12.1); Dysthymia: 17 (13.1) [1] 12.3 (4.8) [2]
Mean months (SD) since onset of current episode	72.3 (98.4) [1] 353.3 (125.9) [2]
No. (SD) of previous depressive episodes	Mean NR (64% ≥1 previous episodes of major depression) [1]

Update 2017

	Sertraline versus imipramine
	NR[2]
Previous treatment	59% psychotherapy; 20% prior adequate trial of antidepressants (defined as at least 150mg/day of amitriptyline or 20mg/day of fluoxetine or their equivalents taken for ≥4 weeks); 43% no previous antidepressant pharmacotherapy[1] NR[2]
Baseline severity	HAMD 25.1 (More severe) [1] HAMD 13.1 (Less severe) [2]
Intervention details	Sertraline
Intervention dose	50-200mg/day (mean final dose 141mg [SD=59.4]) [1] 50-200mg/day (mean final dose 139.6mg [SD=58.5]) [2]
Comparator details (mean dose, if applicable)	Imipramine 50-300mg/day (mean final dose 200.2mg [SD=82.1]) [1] Imipramine 50-300mg/day (mean final dose 198.9mg [SD=91.2]) [2]
Treatment length (weeks)	12
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Keller 1998a; [2]Thase 1996 Thase 1996[2] is a three-armed trial but, where possible, data is extracted for only the two relevant arms here	

Table 168: Summary of findings table for SSRIs compared to other pharmacological interventions

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Imipramine	Corresponding risk Sertraline				
Remission Number of people scoring ≤4 on Hamilton Rating Scale for Depression (HAM-D)/≤7 on HAM-D AND much/very much improved on CGI-I (score 1-2) Follow-up: mean 12 weeks	**Study population** **260 per 1000**	**289 per 1000** (232 to 362)	**RR 1.11** (0.89 to 1.39)	893 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	Moderate **282 per 1000**	**313 per 1000** (251 to 392)				
Response Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND HAMD≤15 AND much/very much improved on CGI-I (score 1-2) AND CGI-S≤3 (mildly	**Study population** **565 per 1000**	**548 per 1000** (486 to 622)	**RR 0.97** (0.86 to 1.1)	893 (2 studies)	⊕⊕⊖⊖ **low**[3,4]	
	Moderate **577 per 1000**	**560 per 1000** (496 to 635)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Imipramine	Sertraline				
ill)/Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 12 weeks						
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 12 weeks		The mean depression symptomatology in the intervention groups was **0.05 standard deviations higher** (0.19 lower to 0.29 higher)		270 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	SMD 0.05 (-0.19 to 0.29)
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 12 weeks	**Study population**		**RR 0.61** (0.39 to 0.95)	905 (2 studies)	⊕⊖⊖⊖ **very low**[3,4,6,7]	
	275 per 1000	**168 per 1000** (107 to 262)				
	Moderate					
	285 per 1000	**174 per 1000** (111 to 271)				
Discontinuation due to adverse events Number of participants discontinuing due to adverse events Follow-up: mean 12 weeks	**Study population**		**RR 0.45** (0.29 to 0.71)	905 (2 studies)	⊕⊖⊖⊖ **very low**[3,4,7]	
	145 per 1000	**65 per 1000** (42 to 103)				
	Moderate					
	152 per 1000	**68 per 1000** (44 to 108)				

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline and method of allocation concealment unclear. Unclear blinding of intervention administration and outcome assessment. Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Funding from pharmaceutical company
[4] Unclear randomisation method and method of allocation concealment. Blinding of intervention administration and outcome assessment is unclear
[5] N<400
[6] I-squared>50%
[7] Events<300

Update 2017

Table 169: Study information table for trials included in the meta-analysis of SSRIs combined with a psychological intervention versus a psychological intervention-only

	Sertraline + IPT versus IPT-only
Total no. of studies (N randomised)	2 (434)
Study ID	Browne 2002[1] Markowitz 2005[2]
Country	Canada[1] US[2]
Chronic definition	Dysthymia[1] Early-onset (<21 years) dysthymic disorder[2]
Age range (mean)	Range NR (42.1) NR by arm (for all four arms of study: Range NR [42.3])
Sex (% female)	NR by arm (for all three arms of study: 68) [1] NR by arm (for all four arms of study: 63) [2]
Ethnicity (% BME)	NR[1] NR by arm (for all four arms of study: 37) [2]
Mean age (SD) at first onset of depression	NR[1] NR (inclusion criteria <21 years) [2]
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	MADRS 25.3 (Less severe) [1] HAMD 19.3 (Less severe) [2]
Intervention details	IPT (followed the manual by Weissman and Klerman 1993 and Klerman et al. 1984) + sertraline[1] IPT for dysthymic disorder (IPT-D; followed manual by Markowitz 1998) + sertraline[2]
Intervention dose	12x 1-hour sessions (mean attended 8.9 sessions [SD=2.6]) + 50-200mg/day of sertraline[1] 16-18 x 50-min sessions (mean attended 12.8 sessions [SD=4.01]) + 50-200mg/day (mean daily dose 116.3 mg/day [SD=43.9]) [2]
Comparator details (mean dose, if applicable)	IPT (followed the manual by Weissman and Klerman 1993 and Klerman et al. 1984). 12x 1-hour sessions (mean attended 10 sessions) [1] IPT for dysthymic disorder (IPT-D; followed manual by Markowitz 1998). 16-18 x 50-min sessions (mean attended 13.2 sessions [SD=4.0])[2]
Treatment length (weeks)	26[1] 16[2]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Browne 2002; [2]Markowitz 2005 Browne 2002 is a three-armed trial and Markowitz 2005 is a four-armed trial but, where possible, data is extracted for only the two relevant arms here	

Update 2017

Table 170: Summary of findings table for SSRIs combined with a psychological intervention compared to a psychological intervention-only

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	IPT-only	Sertraline + IPT				
Remission Number of people scoring <7 on Hamilton Rating Scale for Depression (HAM-D) AND >50% improvement on HAMD AND GAF score>70 Follow-up: mean 16 weeks	**Study population** **217 per 1000**	**524 per 1000** (217 to 1000)	**RR 2.41** (1 to 5.79)	44 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	Moderate **217 per 1000**	**523 per 1000** (217 to 1000)				
Response Number of people showing ≥40% improvement on Montgomery Asberg Depression Rating Scale (MADRS)/≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 16-26 weeks	**Study population** **453 per 1000**	**570 per 1000** (475 to 688)	**RR 1.26** (1.05 to 1.52)	434 (2 studies)	⊕⊖⊖⊖ **very low**[2,4,5]	
	Moderate **407 per 1000**	**513 per 1000** (427 to 619)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 16-26 weeks		The mean depression symptomatology in the intervention groups was **0.5 standard deviations lower** (0.7 to 0.31 lower)		434 (2 studies)	⊕⊖⊖⊖ **very low**[4,5]	SMD -0.5 (-0.7 to -0.31)
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 16 weeks	**Study population** **174 per 1000**	**191 per 1000** (54 to 668)	**RR 1.1** (0.31 to 3.84)	44 (1 study)	⊕⊖⊖⊖ **very low**[1,5,6]	
	Moderate **174 per 1000**	**191 per 1000** (54 to 668)				

Notes:
[1] High risk of bias associated with randomisation method due to significant difference between groups at baseline, and method of allocation concealment is unclear. Non-blind intervention administrator(s) and participants, although outcome assessors are blind
[2] Events<300
[3] Study partially funded by pharmaceutical company
[4] High risk of bias associated with randomisation method due to significant difference between groups

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	IPT-only	Sertraline + IPT				
at baseline. Non-blind intervention administrator(s) and participants, although outcome assessors are blind [5] Data is not reported for all outcomes and funding from pharmaceutical company [6] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)						

Table 171: Study information table for trials included in the meta-analysis of TCAs versus placebo

	TCAs versus placebo
Total no. of studies (N randomised)	7 (1002)
Study ID	Bakish 1993a[1] Boyer 1996 (study 1)[2] Boyer 1996 (study 2)/Lecrubier 1997[3] Kocsis 1988a[4] Stewart 1989/1993[5] Thase 1996[6] Versiani 1997[7]
Country	Canada[1] France[2,3] US[4,5,6] Unclear ('3 countries')[7]
Chronic definition	Dysthymia[1] Dysthymia or double depression[2] Mixed (41% dysthymic disorder; 18% double depression and 40% major depression in partial remission)[3] Double depression (96%; + 4% dysthymic disorder)[4] Dysthymia (sub-analysis of broader depressive disorder sample)[5] Early-onset (<21 years) dysthymia[6] Dysthymia (68%; + 32% double depression)[7]
Age range (mean)	NR[1] Range NR (48.3)[2] 18-73 (43.5)[3] Range NR (39)[4] NR by arm (for all three arms of study: Range NR [37.3])[5] Range NR (41.3)[6] 18-65 (41.5)[7]
Sex (% female)	NR[1] 77[2] 54[3] 70[4] NR by arm (for all three arms of study: 30)[5] 63[6] 72[7]

Update 2017

	TCAs versus placebo
Ethnicity (% BME)	NR[1,2,3,4,7] NR by arm (for all three arms of study: 9) [5] 5[6]
Mean age (SD) at first onset of depression	NR[1,2,3] 20 (13) [4] NR by arm (for all three arms of study: 20.9 [11.8])[5] 12.4 (4.8) [6] NR (36% early onset) [7]
Mean months (SD) since onset of current episode	NR[1,2,3] 228 (192) [4] NR by arm (for all three arms of study: 90.0 [102.7]) [5] 346.3 (128.4) [6] 138.0 (114.0) [7]
No. (SD) of previous depressive episodes	NR
Previous treatment	NR[1,2,3,5,6,7] 71% psychotherapy; 8% adequate trial of TCA; 33% any TCA treatment[4]
Baseline severity	HAMD 15.6 (Less severe) [1] MADRS 17.9 (Less severe) [2] MADRS 25.0 (Less severe) [3] HAMD 22.8 (Less severe) [4] NR by arm (for all three arms of study: HAMD 13.0 [Less severe]) [5] HAMD 13.0 (Less severe) [6] HAMD 20.0 (Less severe) [7]
Intervention details	Imipramine[1,3,4,5,6,7] Amineptine[2]
Intervention dose	50mg/day[1] 200mg/day[2] 50-100mg/day[3] 100-300mg/day[4] ≤300mg/day (mean dose NR for dysthymia subgroup but across broader depression sample: 265mg [SD=47]) [5] 50-300mg/day (mean final dose 198.9mg [SD=91.2]) [6] 25-250mg/day (mean final dose 204mg [SD=64]) [7]
Comparator details (mean dose, if applicable)	Placebo[1,2,3,4,6] Placebo ≤6 tablets (mean dose NR for dysthymia subgroup but across broader depression sample: 5.7 tablets [SD=0.6]) [5] Placebo 2-5 tablets/day (final mean dose 4.5 tablets [SD=1.0]) [7]
Treatment length (weeks)	7[1] 13[2] 26[3] 6[4,5] 12[6] 8[7]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation	

Update 2017

	TCAs versus placebo
[1]Bakish 1993a; [2]Boyer 1996 (study 1); [3]Boyer 1996 (study 2)/Lecrubier 1997; [4]Kocsis 1988a; [5]Stewart 1989/1993; [6]Thase 1996; [7]Versiani 1997 Boyer 1996 (study 1) [2], Boyer 1996 (study 2)/Lecrubier 1997[3], Stewart 1989/1993[5], Thase 1996[6] and Versiani 1997[7] are three-armed trials but, where possible, data is extracted for only the two relevant arms here. Stewart 1989/1993[5] included participants with atypical depression, dysthymic disorder and major depression but data only extracted for the dysthymic disorder subgroup for this review	

1 **Table 172: Summary of findings table for TCAs compared to 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				
	Placebo	TCAs				
Remission (imipramine) Number of people scoring ≤4/6 on Hamilton Rating Scale for Depression (HAM-D)/<8 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: 6-26 weeks	**Study population**		**RR 1.38** (1.02 to 1.86)	667 (4 studies)	⊕⊕⊖⊖ **low**[1,2]	
	239 per 1000	**330 per 1000** (244 to 444)				
	Moderate					
	192 per 1000	**265 per 1000** (196 to 357)				
Response (any TCA) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I)/Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 6-26 weeks	**Study population**		**RR 1.85** (1.51 to 2.26)	831 (5 studies)	⊕⊕⊕⊖ **moderate**[1]	
	361 per 1000	**668 per 1000** (545 to 816)				
	Moderate					
	333 per 1000	**616 per 1000** (503 to 753)				
Response (imipramine) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I)/Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: 6-26 weeks	**Study population**		**RR 1.86** (1.43 to 2.4)	658 (4 studies)	⊕⊕⊖⊖ **low**[1,3]	
	371 per 1000	**690 per 1000** (530 to 890)				
	Moderate					
	338 per 1000	**629 per 1000** (483 to 811)				
Response (amineptine) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I)	**Study population**		**RR 1.92** (1.35 to 2.73)	173 (1 study)	⊕⊕⊖⊖ **low**[2,4]	
	321 per 1000	**617 per 1000** (434 to 877)				
	Moderate					

Update 2017

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	TCAs				
Follow-up: mean 13 weeks	**321 per 1000**	**616 per 1000** (433 to 876)				
Depression symptomatology (any TCA) Hamilton Rating Scale for Depression (HAM-D; change score)/Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: 8-13 weeks		The mean depression symptomatology (any TCA) in the intervention groups was **0.63 standard deviations lower** (0.95 to 0.3 lower)		679 (3 studies)	⊕⊕⊖⊖ **low**[1,3]	SMD -0.63 (-0.95 to -0.3)
Depression symptomatology (imipramine) Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: 8-12 weeks		The mean depression symptomatology (imipramine) in the intervention groups was **0.64 standard deviations lower** (1.21 to 0.08 lower)		467 (2 studies)	⊕⊖⊖⊖ **very low**[1,5]	SMD -0.64 (-1.21 to -0.08)
Depression symptomatology (amineptine) Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: mean 13 weeks		The mean depression symptomatology (amineptine) in the intervention groups was **0.61 standard deviations lower** (0.88 to 0.33 lower)		212 (1 study)	⊕⊕⊖⊖ **low**[4,6]	SMD -0.61 (-0.88 to -0.33)
Discontinuation for any reason (any TCA) Number of participants discontinuing for any reason including adverse events Follow-up: 6-26 weeks	**Study population**		**RR 1.06** (0.85 to 1.31)	935 (6 studies)	⊕⊕⊖⊖ **low**[7,8]	
	289 per 1000	**306 per 1000** (246 to 379)				
	Moderate					
	316 per 1000	**335 per 1000** (269 to 414)				
Discontinuation for any reason (imipramine) Number of participants discontinuing for any reason including adverse	**Study population**		**RR 1.11** (0.83 to 1.49)	716 (5 studies)	⊕⊕⊖⊖ **low**[7,8]	
	259 per 1000	**288 per 1000** (215 to 386)				
	Moderate					

Update 2017

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	TCAs				
events Follow-up: 6-26 weeks	**243 per 1000**	**270 per 1000** (202 to 362)				
Discontinuation for any reason (amineptine) Number of participants discontinuing for any reason including adverse events Follow-up: mean 13 weeks	**Study population**		**RR 0.93** (0.66 to 1.31)	219 (1 study)	⊕⊖⊖⊖ **very low**[9,10]	
	389 per 1000	**362 per 1000** (257 to 509)				
	Moderate					
	389 per 1000	**362 per 1000** (257 to 510)				
Discontinuation due to adverse events (any TCA) Number of participants discontinuing due to adverse events Follow-up: 6-26 weeks	**Study population**		**RR 5.77** (3.09 to 10.79)	935 (6 studies)	⊕⊕⊖⊖ **low**[2,7]	
	21 per 1000	**124 per 1000** (66 to 231)				
	Moderate					
	14 per 1000	**81 per 1000** (43 to 151)				
Discontinuation due to adverse events (imipramine) Number of participants discontinuing due to adverse events Follow-up: 6-26 weeks	**Study population**		**RR 5.87** (3.05 to 11.29)	716 (5 studies)	⊕⊕⊖⊖ **low**[2,7]	
	25 per 1000	**147 per 1000** (76 to 283)				
	Moderate					
	19 per 1000	**112 per 1000** (58 to 215)				
Discontinuation due to adverse events (amineptine) Number of participants discontinuing due to adverse events Follow-up: mean 13 weeks	**Study population**		**RR 4.86** (0.58 to 40.96)	219 (1 study)	⊕⊖⊖⊖ **very low**[9,10]	
	9 per 1000	**45 per 1000** (5 to 379)				
	Moderate					
	9 per 1000	**44 per 1000** (5 to 369)				

Update 2017

Notes:
[1] Unclear (or high risk of bias associated with) randomisation method and unclear method of allocation concealment. Unclear blinding of intervention administration and outcome assessment. Unclear risk of attrition bias (drop-out>20% and/or difference between groups>20% but ITT analysis used)
[2] Events<300
[3] I-squared>50%
[4] Unclear randomisation method and method of allocation concealment. Blinding of intervention administration and outcome assessment is unclear. Unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)

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	TCAs				

[5] I-squared>80%
[6] N<400
[7] Unclear (or high risk of bias associated with) randomisation method and unclear method of allocation concealment. Unclear blinding of intervention administration
[8] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)
[9] Unclear randomisation method and method of allocation concealment. Blinding of intervention administration unclear
[10] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 173: Study information table for trials included in the meta-analysis of TCAs versus other pharmacological interventions

	TCA versus antipsychotic
Total no. of studies (N randomised)	2 (361)
Study ID	Boyer 1996 (study 1)[1] Boyer 1996 (study 2)/Lecrubier 1997[2]
Country	France[1,2]
Chronic definition	Dysthymic disorder or double depression[1] Mixed (40% dysthymic disorder, 19% double depression and 40% major depression in partial remission) [2]
Age range (mean)	Range NR (48.2) [1] 18-73 (42.9) [2]
Sex (% female)	74[1] 52[2]
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	MADRS 17.9 (Less severe) [1] MADRS 24.9 (Less severe) [2]
Intervention details	Amineptine[1] Imipramine[2]
Intervention dose	200mg/day[1] 50-100mg/day[2]
Comparator details (mean dose, if applicable)	Amisulpride 50mg/day
Treatment length (weeks)	13[1] 26[2]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Boyer 1996 (study 1); [2]Boyer 1996 (study 2)/Lecrubier 1997 Boyer 1996 (study 1) and Boyer 1996 (study 2)/Lecrubier 1997 are three-armed trials but, where possible, data is extracted for only the two relevant arms here	

Update 2017

Table 174: Summary of findings table for TCAs compared to other pharmacological 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				
	Antipsychotic	TCA				
Remission (imipramine versus amisulpride) Number of people scoring <8 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 26 weeks	**Study population**		**RR 0.92** (0.59 to 1.45)	146 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	356 per 1000	**328 per 1000** (210 to 516)				
	Moderate					
	356 per 1000	**328 per 1000** (210 to 516)				
Response (any TCA versus amisulpride) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: 13-26 weeks	**Study population**		**RR 0.92** (0.78 to 1.09)	312 (2 studies)	⊕⊕⊖⊖ **low**[4,5]	
	673 per 1000	**619 per 1000** (525 to 734)				
	Moderate					
	673 per 1000	**619 per 1000** (525 to 734)				
Response (amineptine versus amisulpride) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 13 weeks	**Study population**		**RR 0.88** (0.71 to 1.1)	166 (1 study)	⊕⊕⊖⊖ **low**[4,6]	
	701 per 1000	**617 per 1000** (498 to 771)				
	Moderate					
	701 per 1000	**617 per 1000** (498 to 771)				
Response (imipramine versus amisulpride) Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 26 weeks	**Study population**		**RR 0.98** (0.77 to 1.25)	146 (1 study)	⊕⊖⊖⊖ **very low**[1,3,5]	
	644 per 1000	**631 per 1000** (496 to 805)				
	Moderate					
	644 per 1000	**631 per 1000** (496 to 805)				
Depression symptomatology (amineptine versus amisulpride) Montgomery Asberg		The mean depression symptomatology (amineptine versus amisulpride) in the		208 (1 study)	⊕⊕⊖⊖ **low**[4,7]	SMD 0.06 (-0.21 to 0.33)

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Antipsychotic	TCA				
Depression Rating Scale (MADRS; change score) Follow-up: mean 13 weeks		intervention groups was **0.06 standard deviations higher** (0.21 lower to 0.33 higher)				
Discontinuation for any reason (any TCA versus amisulpride) Number of participants discontinuing for any reason including adverse events Follow-up: 13-26 weeks	**Study population**		**RR 1.09** (0.84 to 1.4)	361 (2 studies)	⊕⊕⊖⊖ **low**[8,9]	
	379 per 1000	**413 per 1000** (318 to 530)				
	Moderate					
	383 per 1000	**417 per 1000** (322 to 536)				
Discontinuation for any reason (amineptine versus amisulpride) Number of participants discontinuing for any reason including adverse events Follow-up: mean 13 weeks	**Study population**		**RR 1.01** (0.71 to 1.45)	215 (1 study)	⊕⊖⊖⊖ **very low**[8,10]	
	356 per 1000	**359 per 1000** (253 to 516)				
	Moderate					
	356 per 1000	**360 per 1000** (253 to 516)				
Discontinuation for any reason (imipramine versus amisulpride) Number of participants discontinuing for any reason including adverse events Follow-up: mean 26 weeks	**Study population**		**RR 1.17** (0.81 to 1.68)	146 (1 study)	⊕⊖⊖⊖ **very low**[3,8,9]	
	411 per 1000	**481 per 1000** (333 to 690)				
	Moderate					
	411 per 1000	**481 per 1000** (333 to 690)				
Discontinuation due to adverse events (any TCA versus amisulpride) Number of participants discontinuing due to adverse events Follow-up: 13-26 weeks	**Study population**		**RR 2.16** (1.08 to 4.35)	361 (2 studies)	⊕⊖⊖⊖ **very low**[3,5,8]	
	56 per 1000	**122 per 1000** (61 to 246)				
	Moderate					
	64 per 1000	**138 per 1000** (69 to 278)				

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Antipsychotic	TCA				
Discontinuation due to adverse events (amineptine versus amisulpride) Number of participants discontinuing due to adverse events Follow-up: mean 13 weeks	**Study population**		**RR 2.34** (0.46 to 11.81)	215 (1 study)	⊕⊖⊖⊖ **very low**[8,10]	
	19 per 1000	**45 per 1000** (9 to 227)				
	Moderate					
	19 per 1000	**44 per 1000** (9 to 224)				
Discontinuation due to adverse events (imipramine versus amisulpride) Number of participants discontinuing due to adverse events Follow-up: mean 26 weeks	**Study population**		**RR 2.12** (0.98 to 4.61)	146 (1 study)	⊕⊖⊖⊖ **very low**[3,8,9]	
	110 per 1000	**232 per 1000** (107 to 505)				
	Moderate					
	110 per 1000	**233 per 1000** (108 to 507)				

Notes:
[1] Randomisation method and method of allocation concealment is unclear. Blinding of intervention administrator is also unclear and there is an unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] Data is not reported or cannot be extracted for all outcomes
[4] Randomisation method and method of allocation concealment is unclear. Blinding of intervention administration and outcome assessment is also unclear and there is an unclear risk of attrition bias (drop-out>20% but difference between groups<20% and ITT analysis used)
[5] Events<300
[6] 95% CI crosses both the line of no effect and the threshold for clinically important harm (RR 0.75)
[7] N<400
[8] Randomisation method and method of allocation concealment is unclear. Blinding of intervention administration is also unclear
[9] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)
[10] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 175: Study information table for trials included in the meta-analysis of SNRIs versus placebo

	Duloxetine versus placebo
Total no. of studies (N randomised)	1 (57)
Study ID	Hellerstein 2012
Country	US
Chronic definition	DSM-IV-TR diagnosis of dysthymic disorder or depression NOS
Age range (mean)	19-70 (41.6)

Update 2017

	Duloxetine versus placebo
Sex (% female)	42
Ethnicity (% BME)	30
Mean age (SD) at first onset of depression	19.9 (15)
Mean months (SD) since onset of current episode	95.2 (199.9)
No. (SD) of previous depressive episodes	Mean NR (51% reported no previous major depressive episodes, 21% 1 prior major depression and 28% ≥2 prior episodes of major depression)
Previous treatment	NR
Baseline severity	HAMD 14.5 (Less severe)
Intervention details	Duloxetine
Intervention dose	30-120mg/day (final mean dose 88.97mg [SD=28.33])
Comparator details (mean dose, if applicable)	Placebo 30-120mg/day (final mean dose 100.71mg [SD=27.34])
Treatment length (weeks)	10
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation	

1 **Table 176: Summary of findings table for SNRIs compared to 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				
	Placebo	Duloxetine				
Remission Number of people scoring ≤4 on Hamilton Rating Scale for Depression (HAM-D) AND HAMD item # 1 (depressed mood) score=0 Follow-up: mean 10 weeks	**Study population**		**RR 3.86** (1.47 to 10.13)	57 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	143 per 1000	**551 per 1000** (210 to 1000)				
	Moderate					
	143 per 1000	**552 per 1000** (210 to 1000)				
Response Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) AND much/very much improved on CGI-I (score 1-2) Follow-up: mean 10 weeks	**Study population**		**RR 2.62** (1.31 to 5.24)	57 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	250 per 1000	**655 per 1000** (327 to 1000)				
	Moderate					
	250 per 1000	**655 per 1000** (327 to 1000)				
Depression symptomatology Hamilton Rating Scale		The mean depression symptomatology in		57 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4]	SMD -1.31 (-1.89 to -0.74)

Update 2017

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	Duloxetine				
for Depression (HAM-D; change score) Follow-up: mean 10 weeks		the intervention groups was **1.31 standard deviations lower** (1.89 to 0.74 lower)				

Notes:
[1] High risk of bias associated with randomisation method due to significant group difference at baseline and method of allocation concealment is unclear. Blinding of intervention administration and outcome assessment is also unclear
[2] Events<300
[3] Data cannot be extracted or is not reported for all outcomes and funding from pharmaceutical company
[4] N<400

Table 177: Study information table for trials included in the meta-analysis of MAOIs versus placebo

	Phenelzine versus placebo
Total no. of studies (N randomised)	1 (39)
Study ID	Stewart 1989/1993
Country	US
Chronic definition	Dysthymia (sub-analysis of broader depressive disorder sample)
Age range (mean)	NR by arm (for all three arms of study: Range NR [37.3])
Sex (% female)	NR by arm (for all three arms of study: 30)
Ethnicity (% BME)	NR by arm (for all three arms of study: 9)
Mean age (SD) at first onset of depression	NR by arm (for all three arms of study: 20.9 [11.8])
Mean months (SD) since onset of current episode	NR by arm (for all three arms of study: 90.0 [102.7])
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	NR by arm (for all three arms of study: HAMD 13.0 [Less severe])
Intervention details	Phenelzine
Intervention dose	≤90mg/day (mean dose 73mg [SD=14])
Comparator details (mean dose, if applicable)	Placebo
Treatment length (weeks)	6
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation Stewart 1989/1993 is a three-armed trial but, where possible, data is extracted for only the two relevant arms here. This study also included participants with atypical depression, dysthymic disorder and major depression but data only extracted for the dysthymic disorder subgroup for this review	

Update 2017

Table 178: Summary of findings table for MAOIs compared to 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				
	Placebo	Phenelzine				
Response Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 6 weeks	**Study population**		**RR 1.75** (0.85 to 3.58)	39 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	333 per 1000	**583 per 1000** (283 to 1000)				
	Moderate					
	333 per 1000	**583 per 1000** (283 to 1000)				

Notes:
[1] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administrator(s)
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Data is not reported or cannot be extracted for all outcomes

Table 179: Study information table for trials included in the meta-analysis of MAOIs versus other pharmacological interventions

	Phenelzine versus imipramine
Total no. of studies (N randomised)	2 (69)
Study ID	Stewart 1989/1993[1] Vallejo 1987[2]
Country	US[1] Spain[2]
Chronic definition	Dysthymia (sub-analysis of broader depressive disorder sample)
Age range (mean)	NR by arm (for all three arms of study: Range NR [37.3]) [1] Range NR (40.2) [2]
Sex (% female)	NR by arm (for all three arms of study: 30) [1] 88[2]
Ethnicity (% BME)	NR by arm (for all three arms of study: 9) [1] NR[2]
Mean age (SD) at first onset of depression	NR by arm (for all three arms of study: 20.9 [11.8]) [1] NR[2]
Mean months (SD) since onset of current episode	NR by arm (for all three arms of study: 90.0 [102.7]) [1] 36.6 (4.1) [2]
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	NR by arm (for all three arms of study: HAMD 13.0 [Less severe]) [1] HAMD 20.5 (Less severe) [2]
Intervention details	Phenelzine

Update 2017

	Phenelzine versus imipramine
Intervention dose	≤90mg/day (mean dose 73mg [SD=14]) [1] 30-75mg/day[2]
Comparator details (mean dose, if applicable)	Imipramine ≤300mg/day (mean dose 265mg [SD=47]) [1] Imipramine 100-250mg/day[2]
Treatment length (weeks)	6
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Stewart 1989/1993; [2]Vallejo 1987 Stewart 1989/1993 is a three-armed trial but, where possible, data is extracted for only the two relevant arms here. Stewart 1989/1993 also included participants with atypical depression and major depression and Vallejo 1987 also included participants with major depression with melancholia but data is only extracted for the dysthymic disorder subgroups for this review.	

Table 180: Summary of findings table for MAOIs compared to other pharmacological 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				
	Imipramine	Phenelzine				
Response Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: mean 6 weeks	**Study population**		**RR 0.75** (0.44 to 1.28)	30 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	778 per 1000	**583 per 1000** (342 to 996)				
	Moderate					
	778 per 1000	**584 per 1000** (342 to 996)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D at endpoint) Follow-up: mean 6 weeks		The mean depression symptomatology in the intervention groups was **0.73 standard deviations lower** (1.45 to 0.01 lower)		32 (1 study)	⊕⊕⊖⊖ **low**[4,5]	SMD -0.73 (-1.45 to -0.01)
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 6 weeks	**Study population**		**RR 0.79** (0.2 to 3.07)	39 (1 study)	⊕⊖⊖⊖ **very low**[1,6]	
	200 per 1000	**158 per 1000** (40 to 614)				
	Moderate					
	200 per 1000	**158 per 1000** (40 to 614)				
	Study population					

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Imipramine	Phenelzine				
Discontinuation due to adverse events Number of participants discontinuing due to adverse events Follow-up: mean 6 weeks	**200 per 1000** **Moderate** **200 per 1000**	**158 per 1000** (40 to 614) **158 per 1000** (40 to 614)	**RR 0.79** (0.2 to 3.07)	39 (1 study)	⊕⊖⊖⊖ **very low**[1,6]	

Notes:
[1] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administrator(s)
[2] 95% CI crosses line of no effect and threshold for both clinically important harm (RR 0.75) and clinically important benefit (RR 1.25)
[3] Data is not reported or cannot be extracted for all outcomes
[4] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration and outcome assessment
[5] N<400
[6] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 181: Study information table for trials included in the meta-analysis of RIMAs versus placebo

	Moclobemide versus placebo
Total no. of studies (N randomised)	1 (212)
Study ID	Versiani 1997
Country	Unclear ('3 countries')
Chronic definition	Dysthymia (69%; + 31% double depression)
Age range (mean)	18-65 (40.5)
Sex (% female)	68
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR (34% early onset)
Mean months (SD) since onset of current episode	125.9 (107.9)
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	HAMD 20.5 (Less severe)
Intervention details	Moclobemide
Intervention dose	75-750mg/day (mean final dose 633mg [SD=158])
Comparator details (mean dose, if applicable)	Placebo 2-5 tablets/day (final mean dose 4.5 tablets [SD=1.0])
Treatment length (weeks)	8
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation Versiani 1997 is a three-armed trial but, where possible, data is extracted for only the two relevant arms here.	

Update 2017

Table 182: Summary of findings table for RIMAs compared to 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				
	Placebo	Moclobemide				
Remission Number of people scoring ≤4 on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.92** (1.13 to 3.27)	201 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
	165 per 1000	**317 per 1000** (186 to 539)				
	Moderate					
	165 per 1000	**317 per 1000** (186 to 540)				
Response Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 2.38** (1.71 to 3.31)	201 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
	299 per 1000	**712 per 1000** (511 to 990)				
	Moderate					
	299 per 1000	**712 per 1000** (511 to 990)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 8 weeks		The mean depression symptomatology in the intervention groups was **1.03 standard deviations lower** (1.33 to 0.74 lower)		201 (1 study)	⊕⊕⊖⊖ **low**[1,3]	SMD -1.03 (-1.33 to -0.74)
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 8 weeks	**Study population**		**RR 0.83** (0.42 to 1.67)	212 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	
	144 per 1000	**120 per 1000** (61 to 241)				
	Moderate					
	144 per 1000	**120 per 1000** (60 to 240)				
Discontinuation due to adverse events Number of participants discontinuing due to adverse events Follow-up: mean 8 weeks	**Study population**		**RR 3.37** (0.72 to 15.85)	212 (1 study)	⊕⊖⊖⊖ **very low**[1,4]	
	19 per 1000	**65 per 1000** (14 to 305)				
	Moderate					
	19 per 1000	**64 per 1000** (14 to 301)				

Notes:
[1] Unclear randomisation method and method of allocation concealment, and unclear blinding of

Update 2017

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	Moclobemide				

intervention administration and outcome assessment
[2] Events<300
[3] N<400
[4] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)

Table 183: Study information table for trials included in the meta-analysis of RIMAs versus other pharmacological interventions

	Moclobemide versus imipramine
Total no. of studies (N randomised)	1 (211)
Study ID	Versiani 1997
Country	Unclear ('3 countries')
Chronic definition	Dysthymia (65%; + 35% double depression)
Age range (mean)	18-65 (42.0)
Sex (% female)	72
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR (32% early onset)
Mean months (SD) since onset of current episode	131.7 (114.4)
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	HAMD 20.5 (Less severe)
Intervention details	Moclobemide
Intervention dose	75-750mg/day (mean final dose 633mg [SD=158])
Comparator details (mean dose, if applicable)	Imipramine 25-250mg/day (mean final dose 204mg [SD=64])
Treatment length (weeks)	8
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation Versiani 1997 is a three-armed trial but, where possible, data is extracted for only the two relevant arms here.	

Table 184: Summary of findings table for RIMAs compared to other pharmacological 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				
	Imipramine	Moclobemide				
Remission Number of people scoring ≤4 on Hamilton Rating Scale for	**Study population** **202 per 1000**	**317 per 1000** (194 to 517)	**RR 1.57** (0.96 to 2.56)	198 (1 study)	⊕⊕⊖⊖ **low**[1,2]	

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Imipramine	Moclobemide				
Depression (HAM-D) Follow-up: mean 8 weeks	**Moderate**					
	202 per 1000	**317 per 1000** (194 to 517)				
Response Number of people showing ≥50% improvement on Hamilton Rating Scale for Depression (HAM-D) Follow-up: mean 8 weeks	**Study population**		**RR 1.03** (0.86 to 1.23)	198 (1 study)	⊕⊕⊖⊖ **low**[1,3]	
	691 per 1000	**712 per 1000** (595 to 851)				
	Moderate					
	692 per 1000	**713 per 1000** (595 to 851)				
Depression symptomatology Hamilton Rating Scale for Depression (HAM-D; change score) Follow-up: mean 8 weeks		The mean depression symptomatology in the intervention groups was **0.16 standard deviations lower** (0.44 lower to 0.12 higher)		198 (1 study)	⊕⊕⊖⊖ **low**[1,4]	SMD -0.16 (-0.44 to 0.12)
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: mean 8 weeks	**Study population**		**RR 0.83** (0.41 to 1.65)	211 (1 study)	⊕⊖⊖⊖ **very low**[1,5]	
	146 per 1000	**121 per 1000** (60 to 240)				
	Moderate					
	146 per 1000	**121 per 1000** (60 to 241)				
Discontinuation due to adverse events Number of participants discontinuing due to adverse events Follow-up: mean 8 weeks	**Study population**		**RR 0.61** (0.24 to 1.51)	211 (1 study)	⊕⊖⊖⊖ **very low**[1,5]	
	107 per 1000	**65 per 1000** (26 to 161)				
	Moderate					
	107 per 1000	**65 per 1000** (26 to 162)				

Notes:
[1] Unclear randomisation method and method of allocation concealment, and unclear blinding of intervention administration and outcome assessment
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Events<300
[4] N<400

Update 2017

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Imipramine	Moclobemide				
[5] 95% CI crosses line of no effect and threshold for both clinically important benefit (RR 0.75) and clinically important harm (RR 1.25)						

Table 185: Study information table for trials included in the meta-analysis of antipsychotics versus placebo

	Amisulpride versus placebo
Total no. of studies (N randomised)	2 (358)
Study ID	Boyer 1996 (study 1)[1] Boyer 1996 (study 2)/Lecrubier 1997[2]
Country	France[1,2]
Chronic definition	Dysthymic disorder or double depression[1] Mixed (42% dysthymic disorder, 17% double depression and 41% major depression in partial remission)[2]
Age range (mean)	Range NR (48.0)[1] 18-73 (42.4)[2]
Sex (% female)	73[1] 58[2]
Ethnicity (% BME)	NR
Mean age (SD) at first onset of depression	NR
Mean months (SD) since onset of current episode	NR
No. (SD) of previous depressive episodes	NR
Previous treatment	NR
Baseline severity	MADRS 17.9 (Less severe)[1] MADRS 25.0 (Less severe)[2]
Intervention details	Amisulpride
Intervention dose	50mg/day
Comparator details (mean dose, if applicable)	Placebo
Treatment length (weeks)	13[1] 26[2]
Notes: Abbreviations: mg=milligrams, NR=not reported, SD=standard deviation [1]Boyer 1996 (study 1); [2]Boyer 1996 (study 2)/Lecrubier 1997 Boyer 1996 (study 1) and Boyer 1996 (study 2)/Lecrubier 1997 are three-armed trials but, where possible, data is extracted for only the two relevant arms here.	

Update 2017

Table 186: Summary of findings table for antipsychotics compared to 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				
	Placebo	Amisulpride				
Remission Number of people scoring <8 on Montgomery Asberg Depression Rating Scale (MADRS) Follow-up: mean 26 weeks	**Study population**		**RR 1.62** (0.95 to 2.77)	146 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
	219 per 1000	**355 per 1000** (208 to 607)				
	Moderate					
	219 per 1000	**355 per 1000** (208 to 607)				
Response Number of people rated as much or very much improved on Clinical Global Impressions scale (CGI-I) Follow-up: 13-26 weeks	**Study population**		**RR 2.03** (1.59 to 2.61)	307 (2 studies)	⊕⊕⊖⊖ **low**[1,4]	
	331 per 1000	**672 per 1000** (527 to 864)				
	Moderate					
	332 per 1000	**674 per 1000** (528 to 867)				
Depression symptomatology Montgomery Asberg Depression Rating Scale (MADRS; change score) Follow-up: mean 13 weeks		The mean depression symptomatology in the intervention groups was **0.68 standard deviations lower** (0.97 to 0.4 lower)		206 (1 study)	⊕⊕⊖⊖ **low**[1,5]	SMD -0.68 (-0.97 to -0.4)
Discontinuation for any reason Number of participants discontinuing for any reason including adverse events Follow-up: 13-26 weeks	**Study population**		**RR 0.87** (0.68 to 1.12)	358 (2 studies)	⊕⊕⊖⊖ **low**[6,7]	
	431 per 1000	**375 per 1000** (293 to 483)				
	Moderate					
	441 per 1000	**384 per 1000** (300 to 494)				
Discontinuation due to adverse events Number of participants discontinuing due to adverse events Follow-up: 13-26 weeks	**Study population**		**RR 3.31** (0.92 to 11.9)	358 (2 studies)	⊕⊖⊖⊖ **very low**[3,6,8]	
	17 per 1000	**55 per 1000** (15 to 197)				
	Moderate					
	18 per 1000	**60 per 1000** (17 to 214)				

[1] Unclear randomisation method and method of allocation concealment and unclear blinding of intervention administrator(s). Unclear risk of attrition bias (drop-out>20% but difference between

Update 2017

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	Amisulpride				

groups<20% and ITT analysis used)
[2] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 1.25)
[3] Data is not reported or cannot be extracted for all outcomes
[4] Events<300q
[5] N<400
[6] Unclear randomisation method and method of allocation concealment and unclear blinding of intervention administrator(s).
[7] 95% CI crosses both line of no effect and threshold for clinically important benefit (RR 0.75)
[8] 95% CI crosses both line of no effect and threshold for clinically important harm (RR 1.25)

9.4 Economic evidence

No economic evidence on interventions for adults with chronic depression was identified by the systematic search of the literature. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

9.5 Clinical evidence statements

9.5.1 Psychological interventions

- Very low quality single-RCT evidence (N=125) suggests a clinically important but not statistically significant benefit of problem solving, relative to pill placebo, on the rate of remission in adults with chronic depression.
- Very low quality evidence from 3 RCTs (N=615) suggests neither a clinically important nor statistically significant benefit of a cognitive or cognitive behavioural therapy, relative to an antidepressant, on the rate of remission in adults with chronic depression. While, very low quality evidence from 2 of these RCTs (N=495) suggests a clinically important but not statistically significant benefit in favour of an antidepressant, relative to CBASP, on the rate of response. Very low quality evidence from 2 RCTs (N=458) also suggests a clinically important but not statistically significant benefit in favour of an antidepressant, relative to a cognitive or cognitive behavioural therapy, on depression symptomatology. There was very low quality evidence from 3 RCTs (N=545) suggesting neither clinically important nor statistically significant differences in acceptability or tolerability between a cognitive behavioural therapy relative to an antidepressant, as measured by discontinuation due to any reason.
- Low to very low quality single-RCT evidence (N=436-454) suggests neither a clinically important nor statistically significant benefit of CBASP, relative to nefazodone, on the rate of remission, the rate of response, depression symptomatology or discontinuation due to any reason, in adults with chronic depression. Evidence from this RCT (N=454) did, however, suggest a clinically important and statistically significant benefit of CBASP relative to nefazodone on discontinuation due to adverse events, suggesting greater tolerability of CBASP.
- Very low quality single-RCT evidence (N=59) suggests a clinically important but not statistically significant benefit in favour of escitalopram, relative to CBASP, on the rate of remission and the rate of response in adults with chronic depression. Evidence from this RCT (N=60) suggests a clinically important but not statistically significant effect in the reverse direction on discontinuation due to any reason, with lower drop-out suggesting greater acceptability or tolerability of CBASP relative to escitalopram.

Update 2017

- Very low quality single-RCT evidence (N=120) suggests neither a clinically important nor statistically significant benefit of problem solving, relative to paroxetine, on the rate of remission in adults with chronic depression.
- Low quality single-RCT evidence (N=22) suggests a clinically important and statistically significant benefit in favour of fluoxetine, relative to CBT, on depression symptomatology in adults with chronic depression. Very low quality evidence from this same RCT (N=31) also suggests a clinically important but not statistically significant benefit in favour of fluoxetine on discontinuation due to any reason with higher drop-out in the CBT arm suggesting lower tolerability or acceptability.
- Very low quality evidence from 2 RCTs (N=421) suggests statistically significant benefits, that just miss the threshold for clinical importance, in favour of sertraline relative to IPT on the rate of response and depression symptomatology in adults with chronic depression. While, very low quality evidence from 1 of these RCTs (N=47) suggests a clinically important but not statistically significant benefit in favour of sertraline relative to IPT on the rate of remission. Very low quality single-RCT evidence (N=47) suggests neither a clinically important nor statistically significant difference between IPT and sertraline in acceptability or tolerability, as measured by discontinuation for any reason.
- Very low quality single-RCT evidence (N=50) suggests clinically important and statistically significant benefits in favour of sertraline relative to brief supportive psychotherapy on the rate of remission and depression symptomatology, and a clinically important benefit in favour of sertraline that just misses statistical significance on the rate of response, in adults with chronic depression. Evidence from this same RCT also suggests a clinically important but not statistically significant benefit in favour of sertraline on discontinuation due to any reason with higher drop-out in the brief supportive psychotherapy arm suggesting lower tolerability or acceptability.
- Very low quality evidence from 2 RCTs (N=264-297) suggests clinically important and statistically significant benefits of CBASP, relative to an alternative psychological intervention, on the rate of remission and the rate of response in adults with chronic depression, and a small-to-moderate benefit that just misses statistical significance on depression symptomatology. Very low quality evidence from these 2 RCTs (N=298) also suggests a clinically important but not statistically significant benefit of CBASP relative to an alternative psychological intervention on discontinuation for any reason, with lower drop-out suggesting higher acceptability or tolerability.
- Low quality single-RCT evidence (N=29) suggests a large and statistically significant benefit of CBASP relative to IPT on depression symptomatology in adults with chronic depression, and clinically important benefits that just miss statistical significance on the rate of remission and the rate of response. Very low quality evidence from this same RCT (N=30) suggests neither a clinically important nor statistically significant difference between CBASP and IPT in acceptability or tolerability, as measured by discontinuation for any reason.
- Very low quality single-RCT evidence (N=235-268) suggests a clinically important and statistically significant benefit of CBASP relative to supportive psychotherapy on the rate of response, a small but statistically significant benefit on depression symptomatology, and a clinically important benefit that just misses statistical significance on the rate of remission, in adults with chronic depression. Very low quality evidence from this same RCT (N=268) also suggests a clinically important but not statistically significant benefit of CBASP relative to supportive psychotherapy on discontinuation for any reason, with lower drop-out suggesting higher acceptability or tolerability.
- Very low quality single-RCT evidence (N=49) suggests clinically important but not statistically significant benefits of IPT, relative to brief supportive psychotherapy, on the rate of remission and on acceptability or tolerability (as measured by discontinuation due to any reason) in adults with chronic depression. However, evidence from this same RCT

Update 2017

suggests neither clinically important nor statistically significant benefits of IPT relative to brief supportive psychotherapy on the rate of response or depression symptomatology.

- Very low quality evidence from 2-3 RCTs (N=585-654) suggests clinically important and statistically significant benefits of CBASP combined with treatment as usual or nefazodone, relative to treatment as usual or nefazodone only, on the rate of remission, the rate of response and depression symptomatology in adults with chronic depression. Very low quality evidence from 3 RCTs suggests neither a clinically important nor statistically significant benefit/harm, associated with the addition of CBASP to treatment as usual or nefazodone, on discontinuation due to any reason. However, very low quality single RCT evidence (N=453) did suggest a clinically important and statistically significant benefit of the addition of CBASP to nefazodone on discontinuation due to adverse events, suggesting greater tolerability with the addition of CBASP.
- Very low quality evidence from 2-3 RCTs (N=102-117) suggests clinically important and statistically significant benefits of MBCT combined with treatment as usual, relative to treatment as usual only, on the rate of remission and depression symptomatology in adults with chronic depression. However, low quality evidence from 3 RCTs (N=130) suggests a clinically important but not statistically significant benefit in favour of treatment as usual only on discontinuation due to any reason with higher drop-out in the MBCT arm suggesting lower tolerability or acceptability.
- Very low quality single-RCT evidence (N=88) suggests a large and statistically significant benefit of group CBT combined with treatment as usual, relative to treatment as usual only, on depression symptomatology in adults with chronic depression. Very low quality evidence (N=96) from this same RCT suggests a clinically important benefit that just misses statistical significance on acceptability or tolerability as measured by discontinuation for any reason, with lower drop-out associated with the addition of group CBT.
- Very low quality evidence from 2 RCTs (N=90) suggests a clinically important but not statistically significant benefit of IPT combined with standard pharmacotherapy or sertraline, relative to standard pharmacotherapy or sertraline only, on the rate of remission in adults with chronic depression. However, very low quality evidence from 3-4 RCTs (N=498-522) suggests neither a clinically important nor statistically significant benefit of the addition of IPT to antidepressant treatment on the rate of response or depression symptomatology. Very low quality evidence from 3 RCTs (N=125) suggests neither a clinically important nor statistically significant benefit of the addition of IPT to antidepressant treatment on acceptability or tolerability as measured by discontinuation for any reason.
- Low quality single-RCT evidence (N=45) suggests a clinically important and statistically significant benefit of IPT combined with standard pharmacotherapy relative to standard pharmacotherapy only on the rate of response and depression symptomatology, and a clinically important but not statistically significant benefit on the rate of remission, in adults with chronic depression. However, very low quality evidence from this same RCT (N=45) suggests a clinically important but not statistically significant harm associated with the addition of IPT to standard pharmacotherapy as measured by discontinuation for any reason.
- Very low quality single-RCT evidence (N=45) suggests a clinically important but not statistically significant benefit of IPT combined with sertraline, relative to sertraline only, on the rate of remission in adults with chronic depression. However, very low quality evidence from 2 RCTs (N=453) suggests neither clinically important nor statistically significant benefits of the addition of IPT to sertraline treatment on the rate of response or depression symptomatology. Very low quality single-RCT evidence (N=45) also suggests no significant effects of adding IPT to sertraline on acceptability or tolerability as measured by discontinuation for any reason.

Update 2017

- Very low quality single-RCT evidence (N=24) suggests neither a clinically important nor statistically significant benefit of combined IPT and moclobemide treatment, relative to moclobemide only, on depression symptomatology in adults with chronic depression. However, very low quality evidence from the same RCT (N=35) suggests a clinically important but not statistically significant benefit of the addition of IPT to moclobemide treatment on discontinuation for any reason, with lower drop-out suggesting higher acceptability or tolerability.
- Very low quality single-RCT evidence (N=82) suggests a clinically important and statistically significant benefit of maintenance treatment with CBASP, relative to assessment-only, on preventing relapse and improving depression symptomatology in adults with chronic depression that had partially remitted. Evidence from this same RCT suggests neither a clinically important nor statistically significant difference between maintenance CBASP and assessment-only in acceptability or tolerability, as measured by discontinuation for any reason.
- Very low quality single-RCT evidence (N=33) suggests a clinically important but not statistically significant benefit of Cognitive-Interpersonal Group Psychotherapy for Chronic Depression (CIGP-CD) in combination with fluoxetine, relative to maintenance treatment with fluoxetine-only, on preventing relapse and on acceptability or tolerability (as measured by discontinuation for any reason), in adults with chronic depression that had partially remitted (following acute treatment with fluoxetine). However, evidence from the same RCT (N=35) suggests neither a clinically important nor statistically significant benefit of adding CIGP-CD to fluoxetine treatment on the rate of response.

9.5.23 Pharmacological interventions

Update 2017

- Very low quality evidence from 5-6 RCTs (N=556-578) suggests a clinically important and statistically significant benefit of an SRRI, relative to placebo, on the rate of remission, the rate of response and depression symptomatology, in adults with chronic depression. Very low quality evidence from 6 RCTs (N=593) also suggests a clinically important and statistically significant benefit of SSRIs relative to placebo on discontinuation for any reason. However, very low quality evidence from 4 RCTs (N=385) suggests a clinically important but not statistically significant harm associated with SSRIs as measured by discontinuation due to adverse events.
- Very low quality single-RCT evidence (N=274) suggests a clinically important and statistically significant benefit of sertraline, relative to placebo, on the rate of remission in adults with chronic depression. Very low quality evidence from 2 RCTs (N=339-341) suggests clinically important benefits of sertraline, that just miss statistical significance, on the rate of response and depression symptomatology. Very low quality evidence from both of these RCTs (N=342) also suggests a clinically important and statistically significant benefit of sertraline relative to placebo on discontinuation for any reason. However, very low quality evidence from 1 of these RCTs (N=274) suggests a clinically important but not statistically significant harm associated with sertraline as measured by discontinuation due to adverse events.
- Very low quality evidence from 2 RCTs (N=143) suggests a clinically important and statistically significant benefit of fluoxetine relative to placebo on the rate of response, and a clinically important but not statistically significant benefit on depression symptomatology, in adults with chronic depression. While, very low quality single-RCT evidence (N=111) suggests a clinically important benefit, that just misses statistical significance, on the rate of remission. Very low quality evidence from both of these RCTs (N=175) suggests neither a clinically important nor statistically significant effect of fluoxetine on discontinuation for any reason. However, very low quality evidence from 1 of these RCTs (N=35) suggests a clinically important but not statistically significant harm associated with fluoxetine as measured by discontinuation due to adverse events, although absolute numbers are small.

- Very low quality single-RCT evidence (N=34) suggests a large and statistically significant benefit of escitalopram relative to placebo on depression symptomatology, and clinically important but not statistically significant benefits of escitalopram on the rate of remission and the rate of response, in adults with chronic depression. However, evidence from this same RCT (N=36) suggests a clinically important but not statistically significant harm associated with escitalopram as measured by discontinuation for any reason and discontinuation due to adverse events.
- Low quality single-RCT evidence (N=40) suggests clinically important and statistically significant benefits of paroxetine relative to placebo on the rate of response and depression symptomatology in adults with chronic depression. While, very low quality evidence from 2 RCTs (N=159) suggests a clinically important but not statistically significant benefit of paroxetine on the rate of remission. Low to very low quality evidence from 1 of these RCTs (N=40) also suggests a clinically important but not statistically significant benefit of paroxetine on discontinuation for any reason and none of the participants in this RCT discontinued due to adverse events.
- Moderate quality evidence from 5 RCTs (N=831) suggests a clinically important and statistically significant benefit of a TCA relative to placebo on the rate of response in adults with chronic depression. Low quality evidence from 3 of these RCTs (N=679) also suggests a clinically important and statistically significant benefit of a TCA on depression symptomatology. Low quality evidence from 6 RCTs (N=935) suggests neither a clinically important nor statistically significant effect of a TCA on discontinuation for any reason. However, evidence from these same 6 RCTs suggest a clinically important and statistically significant harm associated with TCAs as measured by discontinuation due to adverse events.
- Low quality evidence from 4 RCTs (N=658-667) suggests clinically important and statistically significant benefits of imipramine, relative to placebo, on the rate of remission and the rate of response in adults with chronic depression. Very low quality evidence from 2 of these RCTs (N=467) suggests a clinically important and statistically significant benefit of imipramine on depression symptomatology. Low quality evidence from 5 RCTs (N=716) suggests neither a clinically important nor statistically significant effect of imipramine on discontinuation for any reason. However, evidence from these same 5 RCTs suggest a clinically important and statistically significant harm associated with imipramine as measured by discontinuation due to adverse events.
- Low quality single-RCT evidence (N=173-212) suggests clinically important and statistically significant benefits of amineptine, relative to placebo, on the rate of response and depression symptomatology in adults with chronic depression. Very low quality evidence from this same RCT (N=219) suggests neither a clinically important nor statistically significant effect of amineptine on discontinuation for any reason, however, evidence from this study does suggest a clinically important (but not statistically significant) harm associated with amineptine as measured by discontinuation due to adverse events.
- Very low quality single-RCT evidence (N=57) suggests clinically important and statistically significant benefits of duloxetine relative to placebo on the rate of remission, the rate of response and depression symptomatology in adults with chronic depression. However, this study did not report discontinuation data so it is not possible to ascertain a proxy for potential harms of duloxetine.
- Very low quality single-RCT evidence (N=39) suggests a clinically important but not statistically significant benefit of phenelzine relative to placebo on the rate of response in adults with chronic depression. However, this study did not report discontinuation data so it is not possible to ascertain a proxy for potential harms of phenelzine.
- Low quality single-RCT evidence (N=201) suggests clinically important and statistically significant benefits of moclobemide relative to placebo on the rate of remission, the rate of response and depression symptomatology in adults with chronic depression. Very low

Update 2017

quality evidence from this same RCT (N=212) suggests neither a clinically important nor statistically significant effect of moclobemide on discontinuation for any reason, however, evidence from this study does suggest a clinically important (but not statistically significant) harm associated with moclobemide as measured by discontinuation due to adverse events.

- Low quality evidence from 2 RCTs (N=307) suggests a clinically important and statistically significant benefit of amisulpride relative to placebo on the rate of response in adults with chronic depression. While, very low quality evidence from 1 of these RCTs (N=146) suggests a clinically important benefit (that just misses statistical significance) of amisulpride on the rate of remission, and low quality evidence from the other RCT (N=206) suggests a clinically important and statistically significant benefit on depression symptomatology. Low to very low quality evidence from both of these RCTs (N=358) suggests neither a clinically important nor statistically significant effect of amisulpride on discontinuation for any reason, however, evidence from these studies does suggest a clinically important harm (that just misses statistical significance) associated with amisulpride as measured by discontinuation due to adverse events.
- Low to very low quality evidence from 2 RCTs (N=893) suggests neither clinically important nor statistically significant benefits of sertraline, relative to imipramine, on the rate of remission or the rate of response in adults with chronic depression. Very low quality evidence from 1 of these RCTs (N=270) also suggests no significant difference between sertraline and imipramine on depression symptomatology. However, very low quality evidence from both of these RCTs (N=905) suggests a clinically important and statistically significant benefit of sertraline relative to imipramine on acceptability or tolerability as measured by discontinuation for any reason or discontinuation due to adverse events.
- Very low quality single-RCT evidence (N=30) suggests a clinically important but not statistically significant benefit in favour of imipramine relative to phenelzine on the rate of response in adults with chronic depression. However, low quality evidence from another RCT (N=32) suggests a clinically important and statistically significant benefit in favour of phenelzine relative to placebo on depression symptomatology. Very low quality evidence from 1 of these RCTs (N=39) suggests no significant differences in acceptability or tolerability between phenelzine and imipramine as measured by discontinuation for any reason and discontinuation due to adverse events.
- Low quality single-RCT evidence (N=198) suggests a clinically important benefit that just misses statistical significance of moclobemide, relative to imipramine, on the rate of remission in adults with chronic depression. However, low to very low quality evidence from this same study (N=198-211) found neither clinically important nor statistically significant differences between moclobemide and placebo on the rate of response, depression symptomatology or discontinuation for any reason. Although evidence from this study does suggest a clinically important but not statistically significant harm of imipramine relative to moclobemide as measured by discontinuation due to adverse events.
- Low quality evidence from 2 RCTs (N=312) suggests neither a clinically important nor statistically significant benefit of a TCA (imipramine or amineptine), relative to an antipsychotic (amisulpride), on the rate of response in adults with chronic depression. Low to very low quality evidence from these same 2 RCTs (N=361) suggests neither a clinically important nor statistically significant difference between a TCA and an antipsychotic on discontinuation for any reason, however, evidence from these studies does suggest a clinically important and statistically significant harm associated with a TCA as measured by discontinuation due to adverse events.
- Low quality single-RCT evidence (N=166-208) suggests neither clinically important nor statistically significant benefits of amineptine, relative to amisulpride, on the rate of response or depression symptomatology in adults with chronic depression. Evidence from this same RCT (N=215) suggests neither a clinically important nor statistically significant

Update 2017

difference between amineptine and amisulpride on discontinuation for any reason, however, evidence from this study does suggest a clinically important (but not statistically significant) harm associated with amineptine as measured by discontinuation due to adverse events.

- Very low quality single-RCT evidence (N=146) suggests neither clinically important nor statistically significant benefits of imipramine, relative to amisulpride, on the rate of remission or the rate of response in adults with chronic depression. Evidence from this same RCT suggests neither a clinically important nor statistically significant difference between imipramine and amisulpride on discontinuation for any reason, however, evidence from this study does suggest a clinically important harm (that just misses statistical significance) associated with imipramine as measured by discontinuation due to adverse events.
- Very low quality evidence from 2 RCTs (N=434) suggests clinically important and statistically significant benefits of sertraline in combination with IPT relative to IPT-only on the rate of response and depression symptomatology in adults with chronic depression. Evidence from 1 of these RCTs (N=44) suggests a clinically important benefit of adding sertraline to IPT, that just misses statistical significance, on the rate of remission. Very low quality evidence from this same RCT (N=44) suggests neither clinically important nor statistically significant effects associated with the addition of sertraline to IPT on acceptability or tolerability as measured by discontinuation for any reason.

9.6 Economic evidence statements

No evidence on the cost effectiveness of interventions for adults with chronic depression is available.

9.7 From evidence to recommendations

9.7.1 Relative values of different outcomes

The GC identified depression symptomology, response, remission, relapse, discontinuation due to adverse events and discontinuation due to any reason (including adverse events) as the critical outcomes for this question.

9.7.2 Trade-off between clinical benefits and harms

Cognitive and cognitive behavioural therapies, in combination with treatment as usual (predominantly psychopharmacology) or a specific antidepressant, appeared consistently to improve depression outcomes for adults with chronic depression compared to psychopharmacological treatment-only. Evidence for improved efficacy with the addition of a psychological intervention to ongoing antidepressant treatment was found for the following specific interventions: MBCT, CBASP and group CBT. However, for MBCT the positive effects on efficacy were considered in the context of the negative effects on the acceptability/tolerability outcome (discontinuation) and the GC decided not to name MBCT as a specific example of an intervention in this class. The GC agreed that the evidence was such that CBASP and CBT should be named as specific examples of interventions in this class but also considered it important to outline some key components that these interventions should include based on the content of the interventions in the evidence reviewed.

The GC noted that although the evidence was in favour of a combined cognitive behavioural and antidepressant treatment, a combined intervention would not be acceptable to everyone. There was consistent low quality evidence for the efficacy of SSRIs alone and evidence on the acceptability and tolerability of SSRIs was better than for other drugs. The GC therefore

Update 2017

agreed that they should recommend SSRIs alone for people with chronic depression who did not wish to receive the psychological component of the combined treatment. There was limited evidence for psychological interventions alone, however, head-to-head comparisons of psychological interventions suggested on the basis of low quality evidence an advantage of CBASP over alternative psychological therapies and the GC therefore agreed that they should recommend considering a cognitive behavioural treatment for people with chronic depression who did not wish to receive the pharmacological intervention component of the combined treatment. A 'consider' rather than 'offer' recommendation was considered appropriate due to the absence of any comparisons of cognitive behavioural treatments-alone against no treatment, treatment as usual, waitlist, or attention-placebo.

The GC considered that although the balance of the evidence was in favour of an SSRI over alternative pharmacological interventions, some people may not be able to tolerate an SSRI or have failed to respond to previous treatment with an SSRI, and for these people an alternative pharmacological intervention would be needed. Given that the majority of the evidence was for first-line treatment of chronic depression and hence recommendations about sequencing represented an extrapolation from the evidence, the GC agreed that it was appropriate to make this a 'consider' rather than an offer recommendation. There was some evidence for benefits of tricyclic antidepressants, moclobemide and amisulpride, and the GC agreed that these should be given as examples of pharmacological interventions that could be considered in circumstances where an SSRI was not appropriate. However, due to concerns around the tolerability of these drugs and potential drug interactions the GC agreed that these should only be prescribed in a specialist setting or after consultation with a specialist.

9.7.3 Trade-off between net health benefits and resource use

Update 2017

The GC considered the high healthcare costs and the burden associated with the presence of chronic depressive symptoms, and the benefits and cost-savings resulting from resolution of chronic depression.

No evidence on the cost-effectiveness of interventions for adults with chronic depression was identified and no further economic analysis was undertaken. The GC considered the results of the economic analysis of treatments of a new depressive episode that was undertaken for the guideline. According to this, for populations with more severe depression, the combination of individual CBT with an antidepressant was likely to be the most cost-effective option for the treatment of new episodes. The GC expressed the view that effective combined treatment with a psychological component that has a focus on chronic depressive symptoms and associated maintaining processes was likely to be cost-effective for people with chronic depression too.

The GC noted that CBASP is not currently in common use in the UK and so there would be some additional costs associated with providing this intervention and training people to use it. However, it was noted that people with chronic depression represent a relatively small proportion of the entire group of people with depression and as such these additional costs were unlikely to be significant. In addition, it was noted that currently there are not many effective treatments available for people with chronic depression and so any increase in costs as a result of these recommendations would likely be balanced by the potential for improved first-line treatment which would reduce the healthcare costs associated with needing to provide one or more further-line treatments.

For people who choose not to have combined treatment, the GC considered SSRIs or cognitive behavioural therapies alone to be alternative cost-effective options, given the results of the guideline economic analyses for the treatment of new episodes, in which SSRIs and psychological interventions were less cost-effective than combined treatment in people with more severe depression, but more cost-effective than clinical management alone.

The GC acknowledged the additional costs associated with provision of antidepressants such as tricyclic antidepressants, moclobemide or amisulpride in specialist settings or after consultation with a specialist. These costs relate to specialist staff time, potentially higher drug acquisition costs (for example, moclobemide and amisulpride, although available in generic form, have higher acquisition costs compared with SSRIs) and costs associated with treatment of side effects. However, the GC considered that these drugs may be the only or best option for a number of people who cannot tolerate an SSRI or have not responded to SSRI treatment, and that, due to their side effect profile, specialist support is needed for safe prescribing and monitoring. Based on the above considerations, the GC made a recommendation for alternative medication, for example tricyclic antidepressants, moclobemide or amisulpride to be considered either in specialist settings or after consultation with a specialist, for people who cannot tolerate an SSRI or have not responded to SSRI treatment.

The GC were mindful that not all people with chronic depression respond to treatment and as a consequence suffer considerable disability and social isolation. They therefore decided to modify the recommendation for this population in the 2009 guideline to offer social or vocational support to people with chronic depression who would benefit from such support. Again given the low numbers to which this would apply and the fact that other non-health agencies may be involved in the provision of these interventions it should not have additional resource implications,

9.7.4 Quality of evidence

The GC noted that all but one outcome had been assessed as either low or very low by GRADE. Most outcomes were downgraded due to imprecision (frequently associated with relatively small sample sizes) and risk of bias. However, the quality of the evidence for interventions for chronic depression was in line with most other areas of the guideline (with the possible exception of the NMA). The results of the evidence for chronic depression were also relatively consistent with interventions that have been found to be effective in other areas of the guideline and this increased the GC's confidence in the results from the evidence.

Update 2017

9.7.5 Other considerations

No evidence was available for psychosocial interventions for chronic depression, as a study on befriending that had been included by the 2009 guideline did not meet our inclusion criteria (different definition of chronic [>1 year] and no mean reported for the duration of depression). However, the GC recognised the potential benefit of additional social or vocational support, particularly given the lack of long-term data on psychological or pharmacological interventions and the potential for poor prognosis and long-term functional impairment, and on this basis the GC agreed to retain the recommendation from the 2009 guideline.

The GC were also aware of the high prevalence of chronic depression in people aged over 75 years and the very limited evidence for the treatment of any type of depression in this age group. They therefore decided to develop a research recommendation to evaluate the effectiveness of psychological, pharmacological or a combination of these interventions in the treatment of adults aged over 75 with chronic depression.

9.8 Recommendations

86. For people with symptoms of chronic depression, consider cognitive behavioural treatments (CBASP and CBT) in combination with antidepressant medication. The cognitive behavioural treatment should:

- have a focus on chronic depressive symptoms
- cover related maintaining processes, for example, avoidance, rumination and interpersonal difficulties. [new 2017]

87. If a person with chronic depression chooses not to have combined treatment, offer:

- an SSRI alone, **or**
- cognitive behavioural treatments (CBASP and CBT) alone. [new 2017]

88. For people with chronic depression who cannot tolerate, or have not responded to, SSRI treatment, consider alternative medication in specialist settings, or after consulting a specialist. Alternatives include:

- tricyclic antidepressants, **or**
- moclobemide, **or**
- amisulpride[s]. [new 2017]

89. For people with chronic depression who have been assessed as likely to benefit from extra social or vocational support, consider:

- befriending in combination with existing antidepressant medication or psychological therapy; this should be done by trained volunteers, typically with at least weekly contact for between 2-6 months
- a rehabilitation programme, if their depression has led to loss of work or their withdrawing from social activities over the longer term. [2017]

90. For people with chronic or treatment-resistant depression who have not responded to the interventions recommended in section 8.8 and 9.8 consider referral to a specialist mental health services for advice and further treatment. [new 2017]

Update 2017

9.9 Research recommendation

3. Are psychological, pharmacological or a combination of these interventions effective and cost effective for the treatment adults aged over 75 with chronic depression?

Statement: A series of randomised controlled trials should be conducted to assess the effectiveness and cost effectiveness of anti-depressants, psychological therapies and the combination of the two in treating people over the age of 75 years with chronic depression. The studies should report on depressive symptoms, personal functioning and quality of life and any adverse events. They should have a follow-up period of at least 12 months.

Rationale: Depression in older people is often not recognised and therefore may go untreated for a significant period of time. The consequences of this are serious as depression, and chronic depression in particular, is associated with an increased risk of developing physical health problems in addition to the burden resulting from the depression. Even when depression is recognised, treatment can be sub-optimal and there is uncertainty about the most effective interventions. Although there are research studies investigating interventions for depression in older adults, many of these study populations have mean

[s] At the time of consultation (July 2017), amisulpride 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 Prescribing guidance: prescribing unlicensed medicines for further information.

ages between 60 and 70 years and the focus is primarily on people with recent onset depression. Randomised controlled trials of psychological, pharmacological or a combination of these interventions in those over 75 with chronic depression are required to assess the relative effectiveness and safety of these interventions.

Update 2017

10 Depression with co-morbidities

10.1 Introduction

10.1.1 Complex depression

Depression associated with other physical (Moussavi, Chatterji et al. 2007) or psychiatric disorders (Kessler, Berglund et al. 2005) is referred to as 'complex depression'. Evidence from the World Health Organization, examining data from 60 different countries in all regions of the world, indicates that depression is significantly more likely in people with chronic physical illness (NICE 2009) and when present significantly worsens the health state associated with those illnesses (Moussavi, Chatterji et al. 2007). Possible reasons for this co-morbidity include: common antecedents, such as childhood adversity increasing both the risk of physical illness and persistent depression (Korkeila, Vahtera et al. 2010, Korkeila, Vahtera et al. 2010, McIntyre, Soczynska et al. 2012); functional and psychological aspects of physical illness leading to new-onset depression (Patten 2001, Ormel, Rijsdijk et al. 2002); or chronic depression leading through biologically plausible mechanisms to new-onset physical illness, including diabetes (Rotella and Mannucci 2013), cardiovascular disease (Kessler and Bromet 2013) and bone disease (Yirmiya and Bab 2009) (Yirmiya and Bab 2009). There is also an established association of chronic depression with additional psychiatric diagnoses, including generalised anxiety disorder (Kessler, Gruber et al. 2008), obsessive-compulsive disorder (Ruscio, Stein et al. 2010), post-traumatic stress disorder (Ginzburg, Ein-Dor et al. 2010), eating disorders (Grilo, White et al. 2009), alcohol use disorders (Kessler, Berglund et al. 2005) and personality disorders (Hirschfeld 1999); and in keeping with the evidence on physical illness, the depression may be primary, secondary or resulting from shared aetiology (Kessler, Gruber et al. 2008). Any number of these problems can present together with a complexity that poses significant challenges to comprehensive formulation and treatment: including clinical uncertainty about how to safely treat depression in the presence of co-morbidity; and the risk that the depression itself is missed (Huffman, Celano et al. 2013). The end result can be under-treatment and worse outcome for both the depression and the associated illness (Gillen, Tennen et al. 2001, Mancuso, Rincon et al. 2001).

The interrelationship between depression and personality disorder (PD) poses particular clinical problems, since both may be viewed as emotion regulation disorders and either may present with irritability, distress or depression at any one time-point. At the outset therefore a careful clinical assessment, including longitudinal assessment of mood, may be needed to make a reliable diagnosis. Additionally, since both depression and PD may share important antecedents, including early trauma, they frequently co-occur (Grant, Chou et al. 2008), so that final diagnosis may conclude an individual has both depression and PD. This reality may sit uncomfortably with separate guidance (for example the NICE guideline on borderline personality disorder: recognition and management CG78 (NICE 2009) and the current guideline) and sometimes separate clinical services for depression and PD. There are associated clinical risks of under-treating, or incorrectly treating, either the PD or the depression. Given all of this particular complexity, the current chapter will focus on the co-occurrence of depression and PD, aiming to give guidance on the available management choices.

10.1.2 Psychotic depression

Psychosis in depression commonly manifests as nihilistic delusions, delusions of guilt, inadequacy and disease, or derogatory auditory hallucinations. People with psychotic depression also demonstrate more severe psychomotor disturbance and greater psychosocial impairment than those without psychosis (Coryell, Leon et al. 1996). In the epidemiologic catchment area study (Johnson, Horwath et al. 1991), 14.7% of patients who

Update 2017

met the criteria for major depression had a history of psychotic features. Limited evidence indicates that psychotic symptoms are more common in samples of older patients than in younger patients (Brodaty, Luscombe et al. 1997). Those with psychotic depression are more likely to require inpatient treatment and to die from suicide or medical causes in the years following their admission (Vythilingam, Chen et al. 2003, Suominen, Haukka et al. 2009). There is some evidence that people with major depression with psychotic features exhibit more frequent relapses or recurrences than patients with non-psychotic depression; however, not all studies are in agreement (Rothschild 2003). Psychotic depression is often not diagnosed accurately, even in specialist settings (Rothschild, Winer et al. 2008), because the psychosis may be subtle, intermittent or concealed. Consequently, it is often inadequately treated (Andreescu, Mulsant et al. 2007).

There has been a long-standing debate as to whether major depression with psychotic features is a distinct syndrome or represents a more severe depressive subtype. The weight of evidence suggests that severity alone does not account for the differences in symptoms, biological features and treatment response (Rothschild 2003, Ostergaard, Bille et al. 2012). Reflecting this, in the DSM-5 classification of mental disorders the presence or absence of psychotic features is a specifier within major depressive disorder, separate from the severity rating. In contrast, in the Tenth Revision of the International Classification of Diseases (ICD 10) (WHO 1992), psychotic depression remains a subtype of severe depression. In recent years, the Psychotic Depression Assessment Scale (PDAS) has been developed for use in the diagnosis of psychotic depression and in the assessment of response to treatment (Park, Choi et al. 2014, Ostergaard, Rothschild et al. 2016). This combines items from the melancholia subscale of the 17-item Hamilton Depression Rating Scale with psychosis items from the Brief Psychiatric Rating Scale.

Possibly germane to the recent DSM-5 reclassification, recent evidence indicates a commonality in brain protein signatures and pathway signalling in psychotic depression and schizophrenia, distinct in both disorders from non-psychotic depression (Martins-de-Souza, Guest et al. 2012, Gottschalk, Wesseling et al. 2014). Although much of this recent interest has been in excitatory neurotransmission (including glutamate signalling), prior work on monoamine transmission also identified relative similarities between depression and schizophrenia (through shared dopaminergic dysfunction) relative to non-psychotic depression. At a treatment level, this work has been supported by the particular importance of antipsychotic (dopamine blocking) drugs in both psychotic depression and schizophrenia (Parker 2012).

The majority of international treatment guidelines on pharmacological approaches to psychotic depression advocate the combination of an antidepressant and antipsychotic medication (Leadholm, Rothschild et al. 2013). However, the use of antidepressant-antipsychotic combinations is associated with potentially serious adverse effects including delayed cardiac conduction, escalating risks of arrhythmia and cardiac arrest. This risk relates to the potential for medication from both drug classes to affect cardiac conduction and can be assessed through measurement of the corrected QT interval on the ECG (Glassman and Bigger 2001). The current evidence base on treatment of psychotic depression will be assessed here.

10.2 Review question

- For adults with complex depression what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination?

The review protocol summary and the eligibility criteria used for this section of the guideline, can be found in Table 187. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Update 2017

Table 187: Clinical review protocol summary for the review of interventions for complex depression

Component	Description
Review question	For adults with complex depression what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination? (RQ2.7)
Population	Adults with complex depression (defined as depression with coexisting personality disorder) Trials included if disaggregated data is available for this population or at least 51% of the participants are eligible for the review
Intervention(s)	Psychological, psychosocial, physical or pharmacological interventions
Comparison	• Treatment as usual • Waitlist • Placebo • Any alternative management strategy
Critical outcomes	• Depression symptomology • Response (e.g. reduction of at least 50% from the baseline score on HAMD/MADRS) • Remission • Relapse • Discontinuation due to side effects • Discontinuation due to any reason (including side effects) Important but not critical outcomes: • Suicide attempts
Study design	RCTs and systematic reviews.

Update 2017

10.2.13 Clinical evidence

70 RCTs from various sources were reviewed at full text for inclusion in this review. These sources included existing systematic reviews (Newton-Howes, Tyrer et al. 2006, Driessen, Cuijpers et al. 2010, Abbass, Town et al. 2011, Town, Abbass et al. 2011, Cuijpers, Sijbrandij et al. 2014) a search of the CENTRAL database, and previous iterations of this guideline (NICE 2004, NICE 2009). Five RCTs were included following review at full text, and these were separated into two comparisons; CBT and behavioural therapies versus psychodynamic therapies, and pharmacotherapy versus combined therapies.

Five RCTs (N =215) met the eligibility criteria for this review: (Liberman and Eckman 1981, Hardy, Barkham et al. 1995, Macaskill and Macaskill 1996, Hellerstein, Rosenthal et al. 1998, Kool, Dekker et al. 2003).

An overview of the trials included in the meta-analyses can be found in Table 188 and Table 189. Further information about both included and excluded studies is contained within Appendix J2.

Summary of findings can be found in Table 190 and Table 191. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

No data were available for the critical outcomes of treatment response or discontinuation due to side effects.

Table 188: Study information table for randomised controlled trials included in the review for CBT and behavioural therapies versus psychodynamic therapies for complex depression

	CBT/behavioural therapies versus psychodynamic therapies
Total no. of studies (N[1])	3 (100)
Study ID	Hardy 1995[2] Hellerstein 1998[3] Liberman 1981[4]
Country	UK[2] USA[3,4]
Depression severity (author description)	Low – high[2] NR[3] Moderate4
Baseline depression score	Low=BDI score of 16-20; moderate=BDI score of 21-26; high=BDI score of 27+2 NR[3] Insight-oriented psychotherapy BDI=26 (15), Behaviour therapy=25 (9)[4]
Personality disorder diagnoses	9 (33%) obtained 2 diagnoses and 18 (67%) obtained 1 - these were distributed amongst obsessive-compulsive, dependent and avoidant types[2] NR[3,4]
Mean age in years	40.3 (9.5)[2], 41.3 (11.1)[3], 29.67 (8.82)[4]
Sex (% female)	53%[2], 55%[3], 67%[4]
Ethnicity (% white)	97%[2], 92%[3], NR[4]
Coexisting conditions/treatments received	22% were on stable regimes of psychotropic medication: hypnotics=5%, anxiolytics=1, hypnotics and anxiolytics=1% antidepressants=15%[2] NR[3,4]
Treatment setting	Outpatient[2] Unclear[3] Inpatient[4]
Treatment length	19-30 weeks[2] NR[3] 10 days[4]
Follow-up length	52 weeks[2] None[3] 2 years[4]
Intervention (mean dose; mg/day)	CBT; 8 sessions across 19 weeks or 16 sessions across 30 weeks[2] Brief supportive psychotherapy; 30-40 sessions[3] Behaviour therapy; 17 hours individual, 10 hours psychodrama and group therapy, 5 hours family therapy[4]
Comparison	Psychodynamic-interpersonal therapy; used Hobson's Conversational Model, 8 sessions across 19 weeks or 16 sessions across 30 weeks[2] Short term Psychodynamic Psychotherapy; 30-40 sessions[3] Insight oriented psychotherapy; 17 hours social skills training, 10 hours anxiety management, 5 hours family negotiation and contingency contracting[4]

Notes:
[1] N=number of patients with complex depression
[2] Hardy 1995
[3] Hellerstein 1998
[4] Liberman 1981

Update 2017

Table 189: Study information table for randomised controlled trials included in the review for pharmacotherapy versus combined therapies for complex depression

	Pharmacotherapy versus combined therapies
Total no. of studies (N[1])	2 (105)
Study ID	Kool 2003[2] Macaskill 1996[3]
Country	Netherlands[2] UK[3]
Depression severity (author description)	NR but excluded patients deemed a suicide risk[2] NR[3]
Baseline depression score	Combi therapy=HAMD-17 score of 20.12 (4.97), Pharm therapy=HAMD-17 score of 20.75 (4.31)[2] NR[3]
Personality disorder diagnoses	30.5% had a paranoid PD, 28.1% avoidant, 29.7% dependent, 27.3% borderline, 8.6% schizoid, 6.2% schizotypal, 5.5% narcissistic, 2.3% antisocial, 1.6% sadistic[2] NR[3]
Mean age in years	NR[2] Pharm group: 37 (12.4), combi group: 39.3 (7.1)[3]
Sex (% female)	62%[2] 70%[3]
Ethnicity (% white)	NR
Coexisting conditions/treatments received	No other treatments received[2] NR[3]
Treatment setting	Outpatient
Treatment length (weeks)	24 weeks
Follow-up length (weeks)	None
Intervention (mean dose; mg/day)	Pharmacotherapy; 3-step model in case of intolerance or lack of efficacy. 1) Fixed dose of 20mg/day fluoxetine 2) 100mg/day amitriptyline rising to 150mg/day and higher if appropriate on basis of plasma concentration 3) 300mg/day moclobemide rising to max 600mg/day.[2] Lofepramine; 35mg 2x daily, increased to 35mg 3x daily from d3, then 70mg 3x daily after d7, increasing to 280mg daily depending upon clinical need and therapeutic response.[3]
Comparison	Combined therapy (pharmacotherapy + short psychodynamic supportive psychotherapy [SPSP]); 3-step model as above plus 16x 45minute sessions of SPSP, 8 sessions weekly then 8 x fortnightly. SPSP focuses on affective, behavioural and cognitive aspects of human relationships using a psychoanalytic frame of reference2 Combined therapy (lofepramine + rational emotive therapy [RET]); pharm. protocol as above plus up to 30x 50 min RET sessions over 24 weeks, with twice weekly sessions permissible in first 5 wks

Notes:
[1] N=number of patients with complex depression.
[2] Kool 2003
[3] Macaskill 1996

Update 2017

Table 190: Summary of findings table for the comparison of CBT/behavioural therapies versus psychodynamic therapies for complex depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Psychodynamic therapies	Risk difference with CBT/behavioural therapies (95% CI)
Depression symptomatology at endpoint BDI	51 (2 studies)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision			The mean depression symptomatology at endpoint in the intervention groups was 6.35 lower (13.18 lower to 0.47 higher)
Depression symptomatology BDI	51 (2 studies) 12 weeks	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision			The mean depression symptomatology in the intervention groups was **0.3 lower** (0.86 lower to 0.25 higher)
Depression symptomatology BDI	24 (1 study) 24 weeks	⊕⊖⊖⊖ **very low**[1,3] due to risk of bias, imprecision			The mean depression symptomatology in the intervention groups was **9.00 lower** (16.09 to 1.91 lower)
Depression symptomatology BDI	24 (1 study) 36 weeks	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision		The mean depression symptomatology in the control groups was 11	The mean depression symptomatology in the intervention groups was **3.00 lower** (11.84 lower to 5.84 higher)
Depression symptomatology BDI	27 (1 study) 1 years	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision		The mean depression symptomatology in the control groups was **12.75**	The mean depression symptomatology in the intervention groups was **0.25 higher** (6.87 lower to 7.37 higher)
Suicide attempts	24 (1 study) 24 weeks	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 0.75** (0.21 to 2.66)	**Study population**	
				333 per 1000	**83 fewer per 1000** (from 263 fewer to 553 more)
				Moderate	
				333 per 1000	**83 fewer per 1000** (from 263 fewer to 553 more)

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Psychodynamic therapies	Risk difference with CBT/behavioural therapies (95% CI)
Suicide attempts (2 year follow-up)	24 (1 study) 2 years	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 0.83** (0.35 to 2.00)	**Study population**	
				500 per 1000	**85 fewer per 1000** (from 325 fewer to 500 more)
				Moderate	
				500 per 1000	**85 fewer per 1000** (from 325 fewer to 500 more)
Discontinuations for any reason	73 (1 study)	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 0.73** (0.33 to 1.60)	**Study population**	
				270 per 1000	**73 fewer per 1000** (from 181 fewer to 162 more)
				Moderate	
				270 per 1000	**73 fewer per 1000** (from 181 fewer to 162 more)

Notes:
[1] High ROB across multiple domains
[2] 95% CI crosses one clinical decision threshold
[3] OIS not met (<400 participants)
[4] 95% CI crosses two clinical decision thresholds

Table 191: Summary of findings table for the comparison of pharmacotherapy versus combined therapies for complex depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Pharmacotherapy versus combi therapy (pharm + SPSP) (95% CI)
Depression symptomatology HAM-D 17	104 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3] due to risk of bias, inconsistency, imprecision			The mean depression symptomatology in the intervention groups was 8 higher (1.35 lower to 17.34 higher)
Depression symptomatology at endpoint (pharm protocol versus	85 (1 study) 24 weeks	⊕⊖⊖⊖ **very low**[4,5] due to risk of bias, imprecision		The mean depression symptomatology at endpoint (pharm protocol versus	The mean depression symptomatology at endpoint (pharm protocol versus pharm + SPSP) in the

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Pharmacotherapy versus combi therapy (pharm + SPSP) (95% CI)
pharm + SPSP) HAM-D 17				pharm + SPSP) in the control groups was **11.1**	intervention groups was **3.79 higher** (0.36 to 7.22 higher)
Depression symptomatology (lofepramine alone versus lofepramine + RET)	19 (1 study)	⊕⊖⊖⊖ **very low**[6,7] due to risk of bias, imprecision		The mean depression symptomatology (lofepramine alone versus lofepramine + ret) in the control groups was **6.7**	The mean depression symptomatology (lofepramine alone versus lofepramine + ret) in the intervention groups was 13.4 higher (5.92 to 20.88 higher)
Remission at endpoint HAM-D 17	85 (1 study) 24 weeks	⊕⊖⊖⊖ **very low**[4,5] due to risk of bias, imprecision	**RR 0.41** (0.2 to 0.86)	**Study population**	
				469 per 1000	**277 fewer per 1000** (from 66 fewer to 376 fewer)
				Moderate	
					-
Discontinuations for any reason	20 (1 study)	⊕⊖⊖⊖ **very low**[3,6] due to risk of bias, imprecision	**RR 0.33** (0.02 to 7.32)	**Study population**	
				100 per 1000	**67 fewer per 1000** (from 98 fewer to 632 more)
				Moderate	
				100 per 1000	**67 fewer per 1000** (from 98 fewer to 632 more)

Notes:
[1] High or unclear ROB across multiple domains
[2] I-squared >80%
[3] 95% CI crosses two clinical decision thresholds
[4] High risk of bias for selective outcome reporting and allocation concealment unlikely to affect results, however unclear effect of bias from missing outcome data
[5] Confidence intervals cross 1 minimally important difference. Sample size less than optimal information size (<400 for continuous outcomes or <300 events for dichotomous outcomes).
[6] High ROB across multiple domains
[7] OIS not met (<400 participants)

Update 2017

1

2

10.2.21 Economic evidence

No economic evidence on interventions for adults with complex depression was identified by the systematic search of the literature. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

10.2.35 Clinical evidence statements

- Very low quality evidence from 3 different RCTs (k=1-2, n=24-73) showed that CBT and behavioural therapies have a clinically important but not statistically significant advantage over psychodynamic therapies on depression symptoms measured with the BDI at endpoint, but that by 1 year follow-up this has not been maintained; that similar numbers of individuals treated with CBT or behavioural therapies and psychodynamic therapies had made suicide attempts at 24 week and 2 year follow-up; and that there was a clinically important but not statistically significant increase in discontinuation rates in those treated with psychodynamic therapies relative to those treated with CBT and behavioural therapies.
- Very low quality evidence from up to 2 RCTs (k=2, n=19-104) showed a clinically important but not statistically significant reduction in depressive symptoms measured on the HAMD-17 in combined therapy overall compared with pharmacotherapy alone, and a significant reduction in those treated specifically with a combination of a pharmacotherapy protocol and a psychodynamic therapy (SPSP) or lofepramine and RET compared with pharmacotherapy or lofepramine alone respectively. Additionally patients treated with combination therapy were more likely to achieve remission than those treated with pharmacotherapy alone, but there was also a clinically important but not statistically significant increase in treatment discontinuations in the combined therapy group.

10.2.44 Economic evidence statements

- No evidence on the cost effectiveness of interventions for adults with complex depression is available.

Update 2017

10.3 From evidence to recommendations

10.3.18 Relative values of different outcomes

The GC identified depression symptomology, response, remission, relapse, discontinuation due to side effects and discontinuation due to any reason (including side effects) as the critical outcomes for this question. However, no data was available for the critical outcomes of response or discontinuation due to side effects. Due to the difficulties engaging this group of patients in treatment and the perception that outcomes may be poorer in this group, when considering the evidence, the GC placed the greatest emphasis on remission data and discontinuation rates.

10.3.26 Trade-off between clinical benefits and harms

The GC noted that this guideline covered people with depression and comorbid personality disorders. The GC were also aware, based on their clinical experience and knowledge, that there was existing NICE guidance about the treatment of people with borderline personality disorders with comorbid depression (CG78) (NICE 2009), which recommended treatment within a well-structured treatment programme for borderline personality disorder. The GC wanted to make recommendations that were in line with this existing NICE guidance. They therefore recommended that referral to a specialist personality treatment disorder programme should be considered.

The GC noted that the greatest evidence for clinical benefit came from studies showing higher remission rates with combined treatment when compared with pharmacological monotherapy.

The GC were also aware, based on their clinical experience and knowledge, of the significant problems in engaging and ensuring uptake in treatment in people with depression comorbid with a personality disorder. They therefore recommended that support should be provided to ensure this happens. A multi-disciplinary setting was considered by the GC to be important due to the complexity of the difficulties experienced by these patients, as this allows access to appropriate expertise. On the basis of their knowledge and clinical experience, and their concerns that some people may not receive an adequate 'dose' of treatment, the GC decided that it was important to specify that it may be necessary to extend the duration of treatment, relative to the length and frequency of treatment that individuals experiencing a depressive episode without a coexisting personality disorder may receive. They noted that this will not always be appropriate, and therefore decided to add the qualifying statement 'when necessary' to indicate that this is best left to clinical judgement.

The GC considered that possible harms would be inadequate duration and intensity of treatment or the provision of ineffective treatment. However they agreed that the percentage of people who were likely to benefit from these recommendations would be higher than those experiencing any harms.

10.3.3 Trade-off between net health benefits and resource use

No evidence on the cost-effectiveness of interventions for adults with complex depression was identified and no further economic analysis was undertaken.

The GC considered that these recommendations would bring practice in line with what has been recommended in CG78 (NICE 2009) and therefore there were unlikely to be any additional costs associated with these recommendations. They also agreed that better treatment of complex depression would probably lead to a reduction in downstream costs associated with not dealing with this condition effectively.

The GC considered the results of the guideline economic analysis on treatment of new episodes of more severe depression, which suggested that combination of antidepressant and high-intensity psychological intervention (CBT) was the most cost-effective treatment among those assessed, and expressed the opinion that, since this treatment showed clinical superiority over pharmacological treatment alone in people with complex depression, it was likely to be cost-effective as well, especially considering the high costs of care associated with sub-optimally treated complex depression, and the cost-savings that would accrue from effective care provided to this population.

10.3.4 Quality of evidence

The quality of the evidence was assessed using GRADE.

The GC noted, based on the evidence, that treatments combining an antidepressant with a high-intensity psychological intervention appeared to be the most effective treatment for people with complex depression. However the GC were mindful that the evidence base for this question was limited in volume, with only five small relevant RCTs identified, and of very low quality for the critical outcomes. Consequently they were only able to recommend combination treatment be 'considered' and they were not able to recommend a specific antidepressant or psychological therapy.

Update 2017

10.4 Recommendations

91. For people with complex depression (depression comorbid with a personality disorder), consider referral to a specialist personality disorder treatment programme. See NICE guidance on borderline personality disorder for recommendations on treatment for personality disorder with coexisting depression. [new 2017]

92. For people with complex depression who have not been able to access, not been helped by or chosen not to be treated in a specialist personality disorder programme, consider a combination of antidepressant medication and CBT. [new 2017]

93. When delivering antidepressant medication and CBT combination treatment for people with complex depression:

- give the person support and encourage them to carry on with the treatment
- provide the treatment in a structured, multidisciplinary setting
- extend the duration of treatment if needed, up to a year. [new 2017]

10.5 Review question

- For adults with psychotic depression what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination?

Update 2017

The review protocol summary and the eligibility criteria used for this section of the guideline, can be found in Table 192. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 192: Clinical review protocol summary for the review of interventions to treat psychotic depression in adults

Component	Description
Review question	For adults with psychotic depression what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination? (RQ2.8)
Population	Adults with psychotic depression (a depressive episode with psychotic features (i.e. delusions and/or hallucinations) in the context of a major depressive disorder)
Intervention(s)	Psychological, psychosocial, physical or pharmacological interventions
Comparison	• Treatment as usual • Waitlist • Placebo • Any alternative management strategy
Critical outcomes	• Depression symptomology • Response (e.g. reduction of at least 50% from the baseline score on HAMD/MADRS) • Remission • Relapse • Discontinuation due to side effects • Discontinuation due to any reason (including side effects)

Component	Description
Study design	RCTs, cluster RCTs and systematic reviews.

10.5.11 Clinical evidence

67 RCTs from various sources were reviewed at full text for inclusion in this review. These sources included an existing systematic review (Wijkstra 2015), a search of the CENTRAL database, and previous iterations of this guideline (2004 and 2009).

Eighteen RCTs met the inclusion criteria in total; fifteen for acute treatment of psychotic depression (McClure, Low et al. 1973, Spiker, Weiss et al. 1985, Spiker and Kupfer 1988, Anton and Burch 1990, Laakman, Faltermaier-Temizel et al. 1995, Bruijn, Moleman et al. 1996, Zanardi, Franchini et al. 1996, Zanardi, Franchini et al. 2000, Mulsant, Sweet et al. 2001, Rothschild, Williamson et al. 2004, van den Broek, Birkenhager et al. 2004, Kunzel, Ackl et al. 2009, Meyers, Flint et al. 2009, Wijkstra, Burger et al. 2010) and three for relapse prevention (Meyers, Klimstra et al. 2001, Navarro, Gasto et al. 2008, Nordenskjöld, Knorring et al. 2013). All studies included in the acute treatment review were pharmacological treatment studies, whilst the included studies in the relapse prevention review consisted of pharmacological and physical (ECT) interventions.

An overview of the trials included in the meta-analyses can be found in Table 193, Table 198, Table 201, Table 204 and Table 208. Further information about both included and excluded studies is contained within Appendix J8.

Summary of findings can be found in Table 194, Table 195, Table 196, Table 197, Table 199, Table 200, Table 202, Table 203, Table 205, Table 206, Table 207, Table 209 and Table 210. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

Update 2017

10.5.1.1 Acute treatment for psychotic depression

10.5.1.1.1 Antidepressant monotherapy versus other pharmacological interventions

Table 193: Study information table for trials included in the meta-analysis of antidepressant monotherapy versus other pharmacological interventions for acute treatment of adults with psychotic depression

	Antidepressants versus placebo	Antidepressants versus antidepressants	Antidepressants versus antipsychotics	Antidepressants versus antipsychotics plus antidepressants
Total no. of studies (N[1])	2 (173)	6 (191)	1 (36)	4 (227)
Study ID	Laakman 1995[2] Spiker 1988[3]	Brujin 1996[4] McClure 1973[5] Wijkstra 2010[6] van den Broek 2004[7] Zanardi 1996[8] Zanardi 2000[9]	Spiker 1985	Anton 1990[10] Kunzel 2008[11] Spiker 1985[12] Wijkstra 2010[6]
Country	Germany[2] USA[3]	Netherlands[4,6,7] Canada[5] Italy[8,9]	USA	USA[10,12] Germany[11] Netherlands[6]
Depression severity	Less severe[2] More severe[3]	More severe[4,5,6,9] Less severe[7] NR[8]	More severe	More severe[10,12,6] NR[11]
Mean age in years	47 (11.4)[2] Amitriptyline: 45.5(13.9), Placebo: 41.3(15.0)[3]	Mirtazapine: 45 (11), Imipramine: 47 (10)[4] 30[5] Imipramine: 52.0(9.6), Venlafaxine: 53.7(6.8)[6] Imipramine: 51(9.1), Fluvoxamine: 53(9.9)[7] Sertraline: 52.6(13.8), Paroxetine: 55.7(13.2)[8] Fluvoxamine: 52.5(9.7), Venlafaxine: 49.0(11.8)[9]	44.1(13.0)	Amoxapine: 44.4 (12.4), combi.: 46.1 (11.5)[10] Trimipramine: 51.4 (12.7), Amitriptyline + haloperidol: 50.6 (13.3)[11] 44.1(13.0)[12] Imipramine: 52.0(9.6), Venlafaxine+Quetiapine: 49.5(11.5)[6]
Sex (% female)	71%[2]	79%[4]	62%	84%[10]

Update 2017

	Antidepressants versus placebo	Antidepressants versus antidepressants	Antidepressants versus antipsychotics	Antidepressants versus antipsychotics plus antidepressants
	62%[3]	50%[5] 47%[6] NR[7] 74%[8] 64%[9]		60%[11] 62%[12] 47%[6]
Ethnicity (% white)	NR	NR[4,5,6,8,9] 68%[7]	93%	71%[10] NR[11,6] 93%[12]
Treatment setting	Outpatient[2] Inpatient[3]	Inpatient	Inpatient	Inpatient[10,12,6] Unclear[11]
Treatment length	6 weeks[2] 4 weeks[3]	4 weeks of predefined blood levels[4,7] 6 weeks[5] 7 weeks[6] 5 weeks[8,9]	4 weeks	4 weeks[10,12] 6 weeks[11] 7 weeks[6]
Intervention (mean dose; mg/day)	Amitriptyline: 50mg b.i.d (max. 200mg, min. 50mg permitted)[2]; 3xdays 50mg, 4xdays 100mg, 7xdays 150mg, 14xdays 200 mg[3]	Imipramine: 37.5-450 mg[4]; plasma levels 200-300µg/L[6]; 150-450 mg daily[7] Clomipramine: 150mg 3x daily[5] Sertraline: d1-3 50mg/day, d4-7 100mg/day, d8 onwards 150mg/day[8] Venlafaxine: 300mg from d8[9]	Amitriptyline: 218mg/day (mean dose)	Amoxapine: 300-500mg/day[10] Trimipramine: 356.1mg/day (mean daily dose)[11] Amitriptyline: 218mg/day (mean dose)[12] Imipramine: plasma levels 200-300µg/L[6]
Comparison	Placebo: 1-2 tablets per day	Mirtazapine: 40-100mg/day[4] Imipramine: 50mg 3x daily[5] Venlafaxine: 375 mg/day[6]	Perphenazine: 50mg/day (mean dose)	Amitriptyline 150-250 mg/day + perphenazine 24-40 mg/day[10]; amitriptyline mean dose 170 mg/day + perphenazine mean dose 54 mg/day[12]

Update 2017

	Antidepressants versus placebo	Antidepressants versus antidepressants	Antidepressants versus antipsychotics	Antidepressants versus antipsychotics plus antidepressants
		Fluvoxamine: 150-1800mg/day[7]; 300mg/day from d8[9] Paroxetine: 50mg/day from d8[8]		Amitriptyline mean dose 184.0 mg/day + haloperidol: 6.3 mg/day[11] Venlafaxine 375mg/day + quetiapine 600 mg/d[6]

Notes:
[1] N=number of patients randomised
[2] Laakman 1995
[3] Spiker 1988
[4] Brujin 1996
[5] McClure 1973
[6] Wijkstra 2010
[7] van den Broek 2004
[8] Zanardi 1996
[9] Zanardi 2000
[10] Anton 1990
[11] Kunzel 2008
[12] Spiker 1985
Note: Mean dose/day and dose ranges/day used in the studies are greater than the maximum doses stated in the SPC for mirtazapine, fluvoxamine, lorazepam, perphenazine (for its licensed indications), alprozolam and amitriptyline. Prescribers should refer to the individual SPCs for doses when prescribing.
Note amoxapine is not available in the UK but is included in the review in order to assess the class effect of pharmacological interventions for depression

Update 2017

1

Table 194: Summary of findings table for antidepressants versus placebo for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus placebo (95% CI)
Depressive symptoms at endpoint (HAMD 17) - TCA versus placebo	136 (1 study)	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision		The mean depressive symptoms at endpoint (HAMD 17) - TCA versus placebo in the control groups was 14.8	The mean depressive symptoms at endpoint (HAMD 17) - TCA versus placebo in the intervention groups was **3 lower** (4.71 to 1.29 lower)
Remission - TCA versus placebo	20 (1 study)	⊕⊖⊖⊖ **very low**[3,4] due to risk of bias, imprecision	**RR 9** (0.55 to 147.95)	**Study population**	
				0 per 1000	**-**
				Moderate	
				0 per 1000	-
Response - TCA versus placebo	136 (1 study)	⊕⊕⊖⊖ **low**[1,5] due to risk of bias, imprecision	Not estimable	See comment	See comment
Discontinuation - TCA versus placebo	173 (2 studies)	⊕⊖⊖⊖ **very low**[3,4] due to risk of bias, imprecision	**RR 1.88** (0.4 to 8.82)	**Study population**	
				34 per 1000	**30 more per 1000** (from 21 fewer to 270 more)
				Moderate	
				115 per 1000	**101 more per 1000** (from 69 fewer to 899 more)

Notes:
[1] Unclear ROB across multiple domains
[2] OIS not met (<400 participants)
[3] High ROB in one domain and unclear in several others
[4] 95% CI crosses two clinical decision thresholds
[5] OIS not met (<300 events)

Update 2017

Table 195: Summary of findings table for antidepressants versus antidepressants for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antidepressant (95% CI)
Depressive symptoms at endpoint - TCA versus SNRI	29 (1 study)	⊕⊕⊖⊖ **low**[1] due to imprecision		The mean depressive symptoms at endpoint - TCA versus SNRI in the control groups was 2.1	The mean depressive symptoms at endpoint - TCA versus SNRI in the intervention groups was **1.1 higher** (1.47 lower to 3.67 higher)
Depressive symptoms at endpoint - TCA (clomipramine) versus TCA (imipramine)	22 (1 study)	⊕⊕⊖⊖ **low**[1] due to imprecision		The mean depressive symptoms at endpoint - TCA (clomipramine) versus TCA (imipramine) in the control groups was **21.3**	The mean depressive symptoms at endpoint - TCA (clomipramine) versus TCA (imipramine) in the intervention groups was **0.3 higher** (8.72 lower to 9.32 higher)
Remission - SSRI versus SNRI	22 (1 study)	⊕⊕⊖⊖ **low**[2,3] due to risk of bias, imprecision	RR 1.5 (0.82 to 2.75)	Study population	
				545 per 1000	**273 more per 1000** (from 98 fewer to 955 more)
				Moderate	
				546 per 1000	**273 more per 1000** (from 98 fewer to 956 more)
Remission - SSRI (sertraline) versus SSRI (paroxetine)	32 (1 study)	⊕⊕⊖⊖ **low**2,3 due to risk of bias, imprecision	**RR 3.37** (1.19 to 9.57)	**Study population**	
				214 per 1000	**508 more per 1000** (from 41 more to 1000 more)
				Moderate	
				214 per 1000	**507 more per 1000** (from 41 more to 1000 more)
Remission - TCA versus SNRI	32 (1 study)	⊕⊕⊕⊖ **moderate**[3] due to imprecision	**RR 0.82** (0.6 to 1.11)	**Study population**	
				917 per 1000	**165 fewer per 1000** (from 367 fewer to 101 more)

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antidepressant (95% CI)
				Moderate	
				917 per 1000	**165 fewer per 1000** (from 367 fewer to 101 more)
Response - TCA versus atypical ADM	30 (1 study)	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 1.29** (0.65 to 2.54)	**Study population**	
				467 per 1000	**135 more per 1000** (from 163 fewer to 719 more)
				Moderate	
				467 per 1000	**135 more per 1000** (from 163 fewer to 719 more)
Response - TCA versus SNRI	33 (1 study)	⊕⊕⊕⊖ **moderate**[3] due to imprecision	**RR 0.87** (0.66 to 1.13)	**Study population**	
				923 per 1000	**120 fewer per 1000** (from 314 fewer to 120 more)
				Moderate	
				923 per 1000	**120 fewer per 1000** (from 314 fewer to 120 more)
Response - TCA versus SSRI	50 (1 study)	⊕⊕⊖⊖ **low**[3,4] due to risk of bias, imprecision	**RR 2.29** (1.14 to 4.58)	**Study population**	
				280 per 1000	**361 more per 1000** (from 39 more to 1000 more)
				Moderate	
				280 per 1000	**361 more per 1000** (from 39 more to 1000 more)
Discontinuation - TCA versus atypical antidepressant	30 (1 study)	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 0.5** (0.19 to 1.31)	**Study population**	
				533 per 1000	**267 fewer per 1000** (from 432 fewer to 165 more)
				Moderate	

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antidepressant (95% CI)
				533 per 1000	**266 fewer per 1000** (from 432 fewer to 165 more)
Discontinuation - TCA versus SSRI	50 (1 study)	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 2** (0.4 to 9.95)	**Study population**	
				80 per 1000	**80 more per 1000** (from 48 fewer to 716 more)
				Moderate	
				80 per 1000	**80 more per 1000** (from 48 fewer to 716 more)
Discontinuation - TCA versus SNRI	33 (1 study)	⊕⊕⊖⊖ **low**[2] due to imprecision	**RR 1.95** (0.23 to 16.79)	**Study population**	
				77 per 1000	**73 more per 1000** (from 59 fewer to 1000 more)
				Moderate	
				77 per 1000	**73 more per 1000** (from 59 fewer to 1000 more)
Discontinuation - TCA (clomipramine) versus TCA (imipramine)	24 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.2** (0.01 to 3.77)	**Study population**	
				167 per 1000	**133 fewer per 1000** (from 165 fewer to 462 more)
				Moderate	
				167 per 1000	**134 fewer per 1000** (from 165 fewer to 463 more)
Discontinuation - SSRI (sertraline) versus SSRI (paroxetine)	32 (1 study)	⊕⊕⊖⊖ **low**[2,3] due to risk of bias, imprecision	**RR 0.07** (0 to 1.2)	**Study population**	
				357 per 1000	**332 fewer per 1000** (from 357 fewer to 71 more)
				Moderate	
				357 per 1000	**332 fewer per 1000** (from 357 fewer to 71 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antidepressant (95% CI)
Discontinuation - SSRI versus SNRI	22 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.2** (0.01 to 3.74)	**Study population**	
				182 per 1000	**145 fewer per 1000** (from 180 fewer to 498 more)
				Moderate	
				182 per 1000	**146 fewer per 1000** (from 180 fewer to 499 more)
Discontinuation due to side effects - TCA (clomipramine) versus TCA (imipramine)	24 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.2** (0.01 to 3.77)	**Study population**	
				167 per 1000	**133 fewer per 1000** (from 165 fewer to 462 more)
				Moderate	
				167 per 1000	**134 fewer per 1000** (from 165 fewer to 463 more)

Notes:
[1] 95% CI crosses two clinical decision thresholds
[2] Unclear ROB across multiple domains
[3] 95% CI crosses one clinical decision threshold
[4] High ROB in at least one domain and unclear in several others

Table 196: Summary of findings table for antidepressants versus antipsychotics for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antipsychotic (95% CI)
Remission - TCA versus antipsychotic	36 (1 study)	⊕⊕⊖⊖ **low**[1] due to imprecision	Not estimable	See comment	See comment
Discontinuation - TCA versus antipsychotic	36 (1 study)	⊕⊕⊖⊖ **low**[1] due to imprecision	Not estimable	See comment	See comment

Notes:
[1] 95% CI crosses two clinical decision thresholds

Table 197: Summary of findings table for antidepressants versus antipsychotics combined with antidepressants for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antipsychotic + antidepressant (95% CI)
Depression symptomatology at endpoint (HAMD-17) - SNRI versus antipsychotic + SNRI	36 (1 study)	⊕⊕⊖⊖ **low**[1] due to imprecision		The mean depression symptomatology at endpoint (HAMD-17) - SNRI versus antipsychotic + SNRI in the control groups was -1.8	The mean depression symptomatology at endpoint (HAMD-17) - SNRI versus antipsychotic + SNRI in the intervention groups was **0.3 lower** (2.44 lower to 1.84 higher)
Depression symptomatology at endpoint (HAMD-17) - Tetracyclic versus antipsychotic +TCA	35 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision		The mean depression symptomatology at endpoint (HAMD-17) - tetracyclic versus antipsychotic +TCA in the control groups was **10.4**	The mean depression symptomatology at endpoint (HAMD-17) - tetracyclic versus antipsychotic +TCA in the intervention groups was **0.9 higher** (5 lower to 6.8 higher)
Depression symptomatology at endpoint (HAMD-17) - TCA versus antipsychotic + SNRI	41 (1 study)	⊕⊕⊖⊖ **low**[1] due to imprecision		The mean depression symptomatology at endpoint (HAMD-17) - TCA versus antipsychotic + SNRI in the control groups was -1.8	The mean depression symptomatology at endpoint (HAMD-17) - TCA versus antipsychotic + SNRI in the intervention groups was 1.4 lower (4.12 lower to 1.32 higher)
Remission - TCA versus TCA + antipsychotic	35 (1 study)	⊕⊕⊕⊖ **moderate**[3] due to imprecision	**RR 0.53** (0.28 to 0.98)	**Study population**	
				778 per 1000	**366 fewer per 1000** (from 16 fewer to 560 fewer)
				Moderate	
				778 per 1000	**366 fewer per 1000** (from 16 fewer to 560 fewer)
				Study population	

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antipsychotic + antidepressant (95% CI)
Remission - SNRI versus antipsychotic + SNRI	36 (1 study)	⊕⊕⊕⊖ **moderate**[3] due to imprecision	**RR 1.1** (0.86 to 1.41)	**833 per 1000**	**83 more per 1000** (from 117 fewer to 342 more)
				Moderate	
				833 per 1000	**83 more per 1000** (from 117 fewer to 342 more)
Remission - TCA versus antipsychotic + SNRI	41 (1 study)	⊕⊕⊕⊖ **moderate**[3] due to imprecision	**RR 1.06** (0.83 to 1.36)	**Study population**	
				833 per 1000	**50 more per 1000** (from 142 fewer to 300 more)
				Moderate	
				833 per 1000	**50 more per 1000** (from 142 fewer to 300 more)
Response - SNRI versus antipsychotic + SNRI	36 (1 study)	⊕⊕⊕⊖ **moderate**[4] due to imprecision	**RR 1.02** (0.88 to 1.18)	**Study population**	
				958 per 1000	**19 more per 1000** (from 115 fewer to 172 more)
				Moderate	
				958 per 1000	**19 more per 1000** (from 115 fewer to 172 more)
Response - Tetracyclic versus antipsychotic + TCA	35 (1 study)	⊕⊖⊖⊖ **very low**[2,3] due to risk of bias, imprecision	**RR 0.75** (0.54 to 1.04)	**Study population**	
				944 per 1000	**236 fewer per 1000** (from 434 fewer to 38 more)
				Moderate	
				944 per 1000	**236 fewer per 1000** (from 434 fewer to 38 more)
Response - TCA versus antipsychotic + SNRI	41 (1 study)	⊕⊕⊕⊖ **moderate**[4] due to imprecision	**RR 0.98** (0.85 to 1.14)	**Study population**	
				958 per 1000	**19 fewer per 1000** (from 144 fewer to 134 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antipsychotic + antidepressant (95% CI)
				Moderate	
				958 per 1000	**19 fewer per 1000** (from 144 fewer to 134 more)
Discontinuation - SNRI versus antipsychotic + SNRI	39 (1 study)	⊕⊕⊖⊖ **low**[1] due to imprecision	**RR 1** (0.1 to 10.04)	**Study population**	
				77 per 1000	**0 fewer per 1000** (from 69 fewer to 695 more)
				Moderate	
				77 per 1000	**0 fewer per 1000** (from 69 fewer to 696 more)
Discontinuation - Tetracyclic versus antipsychotic + TCA	46 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 1.53** (0.69 to 3.4)	**Study population**	
				280 per 1000	**148 more per 1000** (from 87 fewer to 672 more)
				Moderate	
				280 per 1000	**148 more per 1000** (from 87 fewer to 672 more)
Discontinuation - TCA versus antipsychotic + SNRI	46 (1 study)	⊕⊕⊖⊖ **low**[1] due to imprecision	**RR 1.95** (0.36 to 10.58)	**Study population**	
				77 per 1000	**73 more per 1000** (from 49 fewer to 737 more)
				Moderate	
				77 per 1000	**73 more per 1000** (from 49 fewer to 738 more)
Discontinuation - TCA versus antipsychotic + TCA	135 (2 studies)	⊕⊖⊖⊖ **very low**[1,5] due to risk of bias, imprecision	**RR 0.92** (0.51 to 1.66)	**Study population**	
				254 per 1000	**20 fewer per 1000** (from 124 fewer to 167 more)
				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant versus antipsychotic + antidepressant (95% CI)
				235 per 1000	**19 fewer per 1000** (from 115 fewer to 155 more)
Discontinuation due to side effects - TCA versus antipsychotic + TCA	135 (2 studies)	⊕⊖⊖⊖ **very low**[1,5] due to risk of bias, imprecision	**RR 0.52** (0.19 to 1.39)	**Study population**	
				149 per 1000	**72 fewer per 1000** (from 121 fewer to 58 more)
				Moderate	
				134 per 1000	**64 fewer per 1000** (from 109 fewer to 52 more)

Notes:
[1] 95% CI crosses two clinical decision thresholds
[2] High or unclear ROB in most domains
[3] 95% CI crosses one clinical decision threshold
[4] OIS not met (<300 participants)
[5] Unclear ROB across multiple domains

Update 2017

10.5.1.1.2 Combined antidepressant and antipsychotic interventions versus other pharmacological interventions

Table 198: Study information table for trials included in the meta-analysis of combined antidepressant and antipsychotic interventions versus other pharmacological interventions for acute treatment of adults with psychotic depression

	Antidepressants plus antipsychotics versus antidepressants plus placebo	Antidepressants plus antipsychotics versus antipsychotics plus placebo
Total no. of studies (N[1])	1 (36)	1 (259)
Study ID	Mulsant 2001	Meyers 2009
Country	USA	USA
Depression severity	More severe	More severe
Mean age in years	Nortriptyline plus perphenazine=74(8), Nortriptyline plus placebo=71(10)	58.0 (17.7)
Sex (% female)	73%	64%
Ethnicity (% white)	97%	85%
Coexisting conditions/treatments received	NR	NR
Treatment setting	Inpatient	Inpatient or outpatient

	Antidepressants plus antipsychotics versus antidepressants plus placebo	Antidepressants plus antipsychotics versus antipsychotics plus placebo
Treatment length	2-16 weeks	12 weeks
Intervention (mean dose; mg/day)	Nortriptyline 63 mg + perphenazine 19 mg	Olanzapine (minimum target dose 15mg/d) + sertraline (minimum target dose 150mg/d)
Comparison	Nortriptyline 76 mg + placebo	Olanzapine (minimum target dose 15mg/d) + placebo (target dose 150mg/d)
Note: N[1]=number of patients randomised		

Table 199: Summary of findings table for antidepressants plus antipsychotics versus antidepressants combined with placebo for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant + antipsychotic versus antidepressant + placebo (95% CI)
Depression symptomatology at endpoint (HAMD-17) - TCA + antipsychotic versus TCA + placebo	30 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision		The mean depression symptomatology at endpoint (HAMD-17) - TCA + antipsychotic versus TCA + placebo in the control groups was 10.4	The mean depression symptomatology at endpoint (HAMD-17) - TCA + antipsychotic versus TCA + placebo in the intervention groups was **1 higher** (4.24 lower to 6.24 higher)
Remission - TCA + antipsychotic versus TCA + placebo	30 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 1.14** (0.53 to 2.45)	**Study population**	
				438 per 1000	**61 more per 1000** (from 206 fewer to 634 more)
				Moderate	
				438 per 1000	**61 more per 1000** (from 206 fewer to 635 more)
Treatment discontinuation - TCA + antipsychotic versus TCA + placebo	36 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 1.12** (0.26 to 4.81)	**Study population**	
				158 per 1000	**19 more per 1000** (from 117 fewer to 602 more)
				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant + antipsychotic versus antidepressant + placebo (95% CI)
				158 per 1000	**19 more per 1000** (from 117 fewer to 602 more)

Notes:
[1] High ROB in one domain, unclear ROB in several others
[2] 95% CI crosses two clinical decision thresholds

Table 200: Summary of findings table for antidepressants plus antipsychotics versus antipsychotics plus placebo for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antidepressant + antipsychotic versus antipsychotic + placebo (95% CI)
Remission - SSRI + antipsychotic versus antipsychotic + placebo	142 (1 study)	⊕⊕⊕⊖ **moderate**[1] due to imprecision	RR 1.31 (0.98 to 1.75)	**Study population**	
				508 per 1000	**158 more per 1000** (from 10 fewer to 381 more)
				Moderate	
				508 per 1000	**157 more per 1000** (from 10 fewer to 381 more)
Treatment discontinuation - SSRI + antipsychotic versus antipsychotic + placebo	259 (1 study)	⊕⊕⊕⊖ **moderate**[1] due to imprecision	RR 0.7 (0.53 to 0.92)	**Study population**	
				531 per 1000	**159 fewer per 1000** (from 42 fewer to 249 fewer)
				Moderate	
				531 per 1000	**159 fewer per 1000** (from 42 fewer to 250 fewer)

Note:
[1] 95% CI crosses one clinical decision threshold

*10.5.1.1.3*1 *Antipsychotics versus other pharmacological interventions for acute treatment*

2 **Table 201: Study information table for trials included in the meta-analysis of**
3 **antipsychotics versus other pharmacological interventions for acute**
4 **treatment of adults with psychotic depression**

	Antipsychotic versus placebo	Antipsychotic versus antipsychotic plus antidepressant
Total no. of studies (N[1])	2 (201)	1 (73)
Study ID	Rothschild 2004a[2] Rothschild 2004b[3]	Rothschild 2004a
Country	USA	USA
Depression severity	More severe	More severe
Mean age in years	40.7 (12.6)[2] 41.1 (10.4)[3]	40.7 (12.6)
Sex (% female)	52%[2] 50%[3]	52%
Ethnicity (% white)	NR	NR
Coexisting conditions/treatments received	NR	NR
Treatment setting	Inpatient (for at least 1 week) and outpatient	Inpatient (for at least 1 week) and outpatient
Treatment length	8 weeks	8 weeks
Intervention (mean dose; mg/day)	Olanzapine: 5-20mg	Olanzapine: 5-20mg
Comparison	Placebo	Olanzapine plus fluoxetine: 5-20mg olanzapine + 20-80mg fluoxetine
Notes: [1]N=number of patients randomised [2]Rothschild 2004a [3]Rothschild 2004b		

5 **Table 202: Summary of findings table for antipsychotics versus placebo for**
6 **psychotic depression**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antipsychotic versus placebo (95% CI)
Response - Olanzapine versus placebo	116 (2 studies)	⨁⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	RR 0.94 (0.67 to 1.31)	**Study population**	
				528 per 1000	**32 fewer per 1000** (from 174 fewer to 164 more)
				Moderate	
				552 per 1000	**33 fewer per 1000** (from 182 fewer to 171 more)
				Study population	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antipsychotic versus placebo (95% CI)
Treatment discontinuation - Olanzapine versus placebo	201 (2 studies)	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	RR 0.8 (0.58 to 1.09)	**470 per 1000**	**94 fewer per 1000** (from 197 fewer to 42 more)
				Moderate	
				472 per 1000	**94 fewer per 1000** (from 198 fewer to 42 more)

Notes:
[1] Unclear ROB in most domains and high ROB in one
[2] 95% CI crosses two clinical decision thresholds
[3] 95% CI crosses one clinical decision threshold

Table 203: Summary of findings table for antipsychotics versus antipsychotic combined with antidepressant for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antipsychotic versus antipsychotic + antidepressant (95% CI)
Response - antipsychotic versus SSRI + antipsychotic	49 (1 study)	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision	RR 0.45 (0.3 to 0.66)	**Study population**	
				1000 per 1000	**550 fewer per 1000** (from 340 fewer to 700 fewer)
				Moderate	
				1000 per 1000	**550 fewer per 1000** (from 340 fewer to 700 fewer)
Treatment discontinuation - antipsychotic versus antipsychotic +SSRI	73 (1 study)	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	RR 0.62 (0.32 to 1.17)	**Study population**	
				440 per 1000	**167 fewer per 1000** (from 299 fewer to 75 more)
				Moderate	
				440 per 1000	**167 fewer per 1000** (from 299 fewer to 75 more)

Notes:
[1] Unclear ROB in most domains, and high ROB in one
[2] OIS not met (<300 participants)
[3] 95% CI crosses one clinical decision threshold

10.5.1.1.41 Benzodiazepines versus other pharmacological interventions for acute treatment

Table 204: Study information table for trials included in the meta-analysis of benzodiazepines versus other pharmacological interventions for acute treatment of adults with psychotic depression

	Benzodiazepines versus placebo	Benzodiazepines versus antidepressants	Benzodiazepines versus benzodiazepines
Total no. of studies (N[1])	1 (210)	1 (208)	1 (136)
Study ID	Laakman 1995	Laakman 1995	Laakman 1995
Country	Germany	Germany	Germany
Depression severity	Milder depression	Milder depression	Milder depression
Mean age in years	47 (11.4)	47 (11.4)	47 (11.4)
Sex (% female)	71%	71%	71%
Ethnicity (% white)	NR	NR	NR
Coexisting conditions/treatments received	NR	NR	NR
Treatment setting	Outpatient	Outpatient	Outpatient
Treatment length	6 weeks	6 weeks	6 weeks
Intervention (mean dose; mg/day)	Lorazepam: 2.5mg b.i.d (max. of 10mg daily or minimum of 2.5mg permitted) Alprazolam: 1mg b.i.d (max. of 4mg, minimum of 1mg)	Lorazepam: 2.5mg b.i.d (max. of 10mg daily or minimum of 2.5mg permitted) Alprazolam: 1mg b.i.d (max. of 4mg, minimum of 1mg)	Lorazepam: 2.5mg b.i.d (max. of 10mg daily or minimum of 2.5mg permitted)
Comparison	Placebo	Amitriptyline: 50mg b.i.d (max. 200mg, min. 50mg permitted)	Alprazolam: 1mg b.i.d (max. of 4mg, minimum of 1mg)

Notes:

[1]N=number of patients randomised

b.i.d: 2 x daily

Note: Mean dose/day and dose ranges/day used in the studies are greater than the maximum doses stated in the SPC for mirtazapine, fluvoxamine, lorazepam, perphenazine (for its licensed indications), alprozolam and amitriptyline. Prescribers should refer to the individual SPCs for doses when prescribing.

Table 205: Summary of findings table for benzodiazepines versus placebo for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Benzodiazepines versus placebo (95% CI)
Depression symptomatology at endpoint (HAMD-17) - Lorazepam versus placebo	126 (1 study)	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision		The mean depression symptomatology at endpoint (HAMD-17) - lorazepam versus placebo in the	The mean depression symptomatology at endpoint (HAMD-17) - lorazepam versus placebo in the intervention groups was

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Benzodiazepines versus placebo (95% CI)
				control groups was 14.8	**3.7 lower** (5.6 to 1.8 lower)
Depression symptomatology at endpoint (HAMD-17) - Alprazolam versus placebo	129 (1 study)	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision		The mean depression symptomatology at endpoint (HAMD-17) - alprazolam versus placebo in the control groups was **14.8**	The mean depression symptomatology at endpoint (HAMD-17) - alprazolam versus placebo in the intervention groups was **3.2 lower** (5.03 to 1.37 lower)
Response - Lorazepam versus placebo	126 (1 study)	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	RR 3.03 (1.88 to 4.89)	Study population	
				224 per 1000	**454 more per 1000** (from 197 more to 871 more)
				Moderate	
				224 per 1000	**455 more per 1000** (from 197 more to 871 more)
Response - Alprazolam versus placebo	129 (1 study)	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	**RR 2.95** (1.83 to 4.77)	**Study population**	
				224 per 1000	**437 more per 1000** (from 186 more to 844 more)
				Moderate	
				224 per 1000	**437 more per 1000** (from 186 more to 844 more)
Treatment discontinuation - Lorazepam versus placebo	140 (1 study)	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 1.12** (0.42 to 3.03)	**Study population**	
				95 per 1000	**11 more per 1000** (from 55 fewer to 192 more)
				Moderate	
				95 per 1000	**11 more per 1000** (from 55 fewer to 193 more)
Treatment discontinuation - Alprazolam versus placebo	144 (1 study)	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 1.21** (0.46 to 3.16)	**Study population**	
				95 per 1000	**20 more per 1000** (from 51 fewer to 204 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Benzodiazepines versus placebo (95% CI)
				Moderate	
				95 per 1000	**20 more per 1000** (from 51 fewer to 205 more)
Discontinuation due to side effects - Lorazepam versus placebo	140 (1 study)	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 3.36** (0.14 to 81.05)	**Study population**	
				0 per 1000	**-**
				Moderate	
				0 per 1000	**-**
Discontinuation due to side effects - Alprazolam versus placebo	144 (1 study)	⊕⊖⊖⊖ **very low**[1,4] due to risk of bias, imprecision	**RR 7.39** (0.39 to 140.62)	**Study population**	
				0 per 1000	**-**
				Moderate	
				0 per 1000	**-**

Notes:
[1] Unclear ROB in most domains
[2] OIS not met (<400 participants)
[3] OIS not met (<300 events)
[4] 95% CI crosses two clinical decision thresholds

Table 206: Summary of findings table for benzodiazepines versus antidepressants for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Benzodiazepines versus antidepressants (95% CI)
Depression symptomatology at endpoint (HAMD-17) - Lorazepam versus TCA	128 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision		The mean depression symptomatology at endpoint (HAMD-17) - lorazepam versus TCA in the control groups was **11.8**	The mean depression symptomatology at endpoint (HAMD-17) - lorazepam versus TCA in the intervention groups was **0.7 lower** (2.59 lower to 1.19 higher)
Depression symptomatology at endpoint (HAMD-17) - Alprazolam versus TCA	131 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision		The mean depression symptomatology at endpoint (HAMD-17) - alprazolam versus	The mean depression symptomatology at endpoint (HAMD-17) - alprazolam versus TCA in the intervention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Benzodiazepines versus antidepressants (95% CI)
				TCA in the control groups was **11.8**	groups was **0.2 lower** (2.02 lower to 1.62 higher)
Response - Lorazepam versus TCA	128 (1 study)	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	RR 0.88 (0.71 to 1.1)	**Study population**	
				768 per 1000	**92 fewer per 1000** (from 223 fewer to 77 more)
				Moderate	
				768 per 1000	**92 fewer per 1000** (from 223 fewer to 77 more)
Response - Alprazolam versus TCA	131 (1 study)	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	**RR 0.86** (0.69 to 1.07)	**Study population**	
				768 per 1000	**108 fewer per 1000** (from 238 fewer to 54 more)
				Moderate	
				768 per 1000	**108 fewer per 1000** (from 238 fewer to 54 more)
Treatment discontinuation - Lorazepam versus TCA	138 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 2.55** (0.69 to 9.44)	**Study population**	
				42 per 1000	**65 more per 1000** (from 13 fewer to 352 more)
				Moderate	
				42 per 1000	**65 more per 1000** (from 13 fewer to 354 more)
Treatment discontinuation - Alprazolam versus TCA	142 (1 study)	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	**RR 2.74** (0.76 to 9.92)	**Study population**	
				42 per 1000	**73 more per 1000** (from 10 fewer to 372 more)
				Moderate	
				42 per 1000	**73 more per 1000** (from 10 fewer to 375 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Benzodiazepines versus antidepressants (95% CI)
Discontinuation due to side effects - Lorazepam versus TCA	138 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 3.27** (0.14 to 78.87)	**Study population**	
				0 per 1000	**-**
				Moderate	
				0 per 1000	**-**
Discontinuation due to side effects - Alprazolam versus TCA	142 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 7.2** (0.38 to 136.84)	**Study population**	
				0 per 1000	**-**
				Moderate	
				0 per 1000	**-**

Notes:
[1] Unclear ROB in most domains
[2] 95% CI crosses two clinical decision thresholds
[3] 95% CI crosses one clinical decision threshold

Table 207: Summary of findings table for benzodiazepines versus benzodiazepines for psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Benzodiazepines versus benzodiazepines (95% CI)
Depression symptomatology at endpoint (HAMD-17) - Lorazepam versus alprazolam	121 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision		The mean depression symptomatology at endpoint (HAMD-17) - lorazepam versus alprazolam in the control groups was 11.6	The mean depression symptomatology at endpoint (HAMD-17) - lorazepam versus alprazolam in the intervention groups was **0.5 lower** (2.5 lower to 1.5 higher)
Response - Lorazepam versus alprazolam	121 (1 study)	⊕⊕⊖⊖ **low**[1,3] due to risk of bias, imprecision	**RR 1.03** (0.8 to 1.32)	**Study population**	
				661 per 1000	**20 more per 1000** (from 132 fewer to 212 more)
				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Benzodiazepines versus benzodiazepines (95% CI)
				661 per 1000	**20 more per 1000** (from 132 fewer to 212 more)
Treatment discontinuation - Lorazepam versus alprazolam	136 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.93** (0.36 to 2.42)	**Study population**	
				114 per 1000	**8 fewer per 1000** (from 73 fewer to 162 more)
				Moderate	
				114 per 1000	**8 fewer per 1000** (from 73 fewer to 162 more)
Discontinuation due to side effects - Lorazepam versus alprazolam	136 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.35** (0.04 to 3.31)	**Study population**	
				43 per 1000	**28 fewer per 1000** (from 41 fewer to 99 more)
				Moderate	
				43 per 1000	**28 fewer per 1000** (from 41 fewer to 99 more)

Notes:
[1] Unclear ROB across most domains
[2] 95% CI crosses two clinical decision thresholds
[3] 95% CI crosses one clinical decision threshold

10.5.1.21 Relapse prevention for psychotic depression

Table 208: Study information table for trials included in the meta-analysis of interventions for relapse prevention in adults with psychotic depression

	ECT plus antidepressants versus antidepressants (+/- Lithium)	Antidepressants plus antipsychotics versus antidepressants combined with placebo
Total no. of studies (N[1])	2 (54)	1 (28)
Study ID	Navarro 2008[3] Nordenskjold 2013[4]	Meyers 2001
Country	Spain[2] Sweden[3]	USA
Baseline depression severity	More severe	More severe

	ECT plus antidepressants versus antidepressants (+/- Lithium)	Antidepressants plus antipsychotics versus antidepressants combined with placebo
Mean age in years	Nortriptyline: 70.7 (3.4), ECT/Nortriptyline= 70.4 (3.2)[2] ECT Plus Pharmacotherapy=52 (17), Pharmacotherapy Alone= 62 (13)[3]	71.8 (8.4)
Sex (% female)	36%[2] 50%[3]	68%
Ethnicity (% white)	NR	NR
Coexisting conditions/treatments received	NR	NR
Treatment setting	Inpatient and outpatient[2] Inpatient[3]	Inpatient
Acute treatment	ECT	Uncontrolled inpatient treatment
Relapse prevention treatment length	Up to 2 years[2] 1 year[3]	6 months
Relapse prevention intervention (mean dose; mg/day)	Continuation Nortriptyline+ ECT: weekly ECT for 1 month, fortnightly for next month, then monthly. Nortriptyline treatment based upon plasma concentrations[2] Continuation ECT plus pharmacotherapy: unilateral ultrabrief ECT (29x in 1 year), venlafaxine +/- Lithium[3]	Nortriptyline + antipsychotic: 25mg/day on days 1-3, 50mg/day on days 2/3 – 7, dose adjusted for plasma concentration of 50ng/ml-150ng/ml. If nortriptyline contraindicated sertraline 50-100mg/day given. Perphenazine 4mg added at d7, dose titrated over 2 weeks to 120-160mg/day.
Comparison	Continuation Nortriptyline[2] Pharmacotherapy alone (venlafaxine first choice, lithium augmentation offered to all)[3]	Nortriptyline + placebo:25mg/day on days 1-3, 50mg/day on days 2/3 – 7, dose adjusted for plasma concentration of 50ng/ml-150ng/ml. If nortriptyline contraindicated sertraline 50-100mg/day given. Placebo added at d14.

Notes:
[1] N=number of patients randomised
[2] Navarro 2008,
[3] Nordenskold 2013

Table 209: Summary of findings table for ECT plus antidepressants versus antidepressants (+/- lithium) for relapse prevention in psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with ECT + ADM versus ADM (+/- Li) (95% CI)
Relapses	54 (2 studies)	⨁⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	RR 0.65 (0.22 to 1.91)	**Study population**	
				222 per 1000	**78 fewer per 1000** (from 173 fewer to 202 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with ECT + ADM versus ADM (+/- Li) (95% CI)
				Moderate	
				259 per 1000	**91 fewer per 1000** (from 202 fewer to 236 more)
Relapses - ECT + TCA versus TCA	33 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	RR 0.53 (0.05 to 5.31)	**Study population**	
				118 per 1000	**55 fewer per 1000** (from 112 fewer to 507 more)
				Moderate	
				118 per 1000	**55 fewer per 1000** (from 112 fewer to 509 more)
Relapses - ECT + ADM versus ADM (+/- Li augmentation)	21 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.68** (0.2 to 2.33)	**Study population**	
				400 per 1000	**128 fewer per 1000** (from 320 fewer to 532 more)
				Moderate	
				400 per 1000	**128 fewer per 1000** (from 320 fewer to 532 more)

Notes:
[1] High ROB in one domain and unclear in several others
[2] 95% CI crosses two clinical decision thresholds

Table 210: Summary of findings table for antidepressants plus antipsychotics versus antidepressants combined with placebo for relapse prevention in psychotic depression

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with ADM + antipsychotic versus ADM + placebo (95% CI)
Relapses - TCA + antipsychotic versus TCA + placebo	28 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	RR 2.17 (0.5 to 9.35)	**Study population**	
				154 per 1000	**180 more per 1000** (from 77 fewer to 1000 more)
				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with ADM + antipsychotic versus ADM + placebo (95% CI)
				154 per 1000	**180 more per 1000** (from 77 fewer to 1000 more)

Notes:
[1] Unclear ROB in most domains
[2] 95% CI crosses two clinical decision thresholds

10.5.21 Economic evidence

No economic evidence on interventions for adults with psychotic depression was identified by the systematic search of the literature. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

10.5.35 Clinical evidence statements

10.5.3.16 Acute treatment for psychotic depression

10.5.3.1.17 Antidepressant monotherapy versus other pharmacological interventions

- Low-very low quality evidence from up to 2 RCTs (k=1-2, n=20-173) showed lower levels of depressive symptoms, a greater likelihood of remission and response and a clinically important but not statistically significant increase in treatment discontinuation rates at treatment endpoint in patients treated with a TCA than those treated with placebo.
- Low quality evidence from 2 different RCTs (k=1-1, n=22-29) showed no difference in depressive symptoms between patients treated with a TCA and an SNRI, or between those treated with one TCA (clomipramine) and another TCA (imipramine).
- Moderate-low quality evidence from 3 RCTs (k=1-1, n=22-32) showed a clinically important but not statistically significant increase in remission rates in patients treated with an SNRI compared with an SSRI, but no difference between patients treated with a TCA or SNRI, or in patients treated with paroxetine when compared with sertraline.
- Moderate-very low quality evidence from 1 RCT (k=1, n=30) showed a clinically important but not statistically significant increase in response rates in patients treated with atypical antidepressants compared with TCAs, no difference between those treated with a TCA or an SNRI, and greater response rates in those treated with an SSRI compared with a TCA.
- Low-very low quality evidence from 6 RCTs (k=1-1, n=22-5030) showed a clinically important but not statistically significant increase in discontinuation rates in patients treated with an atypical antidepressant compared with a TCA, in those treated with a TCA compared with an SSRI or SNRI, in those treated with a specific TCA (imipramine) compared with another (clomipramine), in those treated with an SNRI compared with an SSRI and in those treated with one SSRI (paroxetine) compared with another (sertraline). There was also a clinically important but not statistically significant increase in discontinuations due to side effects in those treated with one TCA (imipramine) compared with another (clomipramine).
- Low quality evidence from 1 RCT (k=1, n=36) showed a clinically important but not statistically significant increase in rates of remission and discontinuation in patients treated with a TCA compared with an antipsychotic.
- Moderate-very low quality evidence from 2 RCTs (k=1-1, n=35-41) showed no difference in depressive symptoms or response rates between patients treated with an SNRI or TCA

Update 2017

alone and those treated with a combination of SNRI and antipsychotic, or with a tetracyclic antidepressant alone and those treated with a combination of TCA and antipsychotic.

- Moderate quality evidence from 2 RCTs (k=1-1, n=35-41) showed higher remission rates in those patients treated with a combination of TCA and antipsychotic medications compared with a TCA alone, but no difference between those patients treated with a combination of SNRI and an antipsychotic and those treated with an SNRI or TCA alone.
- Moderate quality evidence from 1 RCT (k=1, n=41) showed no difference in response rates between those patients treated with a combination of an SNRI and an antipsychotic and those treated with a TCA alone.
- Low-very low quality evidence from 5 RCTs (k=1-2, n=39-135) showed no difference in discontinuation rates between those patients treated with a combination of an SNRI and an antipsychotic and those treated with an SNRI alone, or those treated with a combination of a TCA and an antipsychotic and those treated with a TCA alone, but a clinically important but not statistically significant increase in discontinuation rates in those patients treated with the combination of a TCA and antipsychotic compared with a tetracyclic alone or a combination of an SNRI and antipsychotic compared with a TCA alone. However there was a clinically important but not statistically significant increase in discontinuation rates due to side effects in patients treated with a TCA alone compared with those treated with a combination of a TCA and an antipsychotic.

10.5.3.1.2 Combined antidepressant and antipsychotic interventions versus other pharmacological interventions

- Very low quality evidence from 1 RCT (k=1-1, n=30-36) showed no difference in depressive symptoms, remission rates or discontinuation rates between patients treated with a TCA combined with an antipsychotic, and those treated with a TCA and placebo pills.
- Moderate quality evidence from 1 RCT (k=1-1, n=142-259) showed a clinically important but not statistically significant increase in remission rates and fewer treatment discontinuations in patients treated with an SSRI plus an antipsychotic when compared with those who were treated with an antipsychotic combined with a placebo.

10.5.3.1.3 Antipsychotic monotherapy versus other pharmacological interventions

- Low-very low quality evidence from 2 RCTs (k=2-2, n=116-201) found no difference in clinical response or discontinuation rates between patients treated with olanzapine or placebo.
- Low quality evidence from 1 RCT (k=1-1, n=49-73) found a higher rate of response and a clinically important but not statistically significant increase in treatment discontinuations in patients treated with a combination of antipsychotic medication and an SSRI when compared with those treated only with an antipsychotic.

10.5.3.1.4 Benzodiazepines versus other pharmacological interventions

- Low-very low quality evidence from one, 3-armed RCT (k=1-1, n=121-129), suggests that both lorazepam and alprazolam are more effective than placebo at reducing depressive symptoms, inducing clinical response by treatment endpoint and that there is no difference between the benzodiazepines and placebo in terms of treatment discontinuation rates, but that there is a clinically important but not statistically significant increase in discontinuation due to side effects when treated with benzodiazepines.
- Low-very low quality evidence from 1 RCT (k=1-1, n=128-131) showed a clinically important but not statistically significant decrease in depression symptoms and increase in treatment discontinuation rates both for any reason and due to side effects of both lorazepam and alprazolam over a tricyclic antidepressant, but no difference in clinical response between benzodiazepines and TCAs.
- Low-very low quality evidence from 1 RCT (k=1-1, n=121-136) demonstrated no difference in depressive symptoms, clinical response or treatment discontinuation rates at

Update 2017

endpoint between patients treated with lorazepam and alprazolam, but a a clinically important but not statistically significant increase in discontinuation due to side effects in patients treated with alprazolam versus lorazepam.

10.5.3.2 Relapse prevention for psychotic depression

- Very low quality evidence from 3 different RCTs (k=1-2, n=21-54) suggests there is a clinically important, but not statistically significant, benefit of receiving a combination of ECT and an antidepressant, including a tricyclic depressant, rather than an antidepressant (with or without lithium augmentation) alone, or from supplementing a tricyclic antidepressant with placebo rather than an antipsychotic, for relapse prevention.

10.5.4 Economic evidence statements

- No evidence on the cost effectiveness of interventions for adults with psychotic depression is available.

10.5.5 From evidence to recommendations

10.5.5.1 Relative values of different outcomes

The GC identified depression symptomology, response, remission, relapse, discontinuation and discontinuation due to side effects to be the critical outcomes for this question. Data were available for all of these critical outcomes.

10.5.5.2 Trade-off between clinical benefits and harms

The greatest evidence of clinical benefit was seen in the RCTs examining the effectiveness of TCAs, the provision of benzodiazepines and augmentation of an antidepressant with an antipsychotic.

The GC noted that TCAs, although highly clinically effective, were associated with higher discontinuation rates in the RCTs as well as having significant cardiovascular risks associated with their use. The evidence for benzodiazepines meanwhile came from a single study and showed greater effectiveness but increased discontinuations due to side effects. Therefore they did not recommend these interventions.

The GC noted that there was little evidence on the use of ECT, and this was not statistically significant. Therefore they decided not to make a recommendation for this intervention. The GC agreed that the evidence for combined treatment with an antidepressant and an antipsychotic presented a moderately consistent picture of clinical benefit and therefore recommended this.

The GC discussed whether patients with psychotic depression could be safely and effectively cared for within primary care services, but judged that their needs would be better met within secondary care services. They specifically discussed whether GPs would be comfortable commencing prescriptions for antipsychotics to augment antidepressant treatment. The GC agreed, based on their knowledge and experience, that this would often not be the case. Consequently they recommended that coordinated multi-professional care would be necessary and people should be referred to specialist mental health services so that the complex needs of this patient group could be dealt with effectively.

The GC were aware that no evidence on psychological interventions for people with psychotic depression had been identified. Based on their knowledge and experience of the use of psychological interventions in the treatment of psychosis, the GC noted that psychological interventions may also be effective for psychotic depression. They therefore agreed that psychological interventions should be reviewed as part of the coordinated multi-professional programme of care in case they were of benefit to the individual.

Update 2017

The GC considered the greatest possible harms to be unacceptable levels of side effects associated with pharmacological treatment and the provision of ineffective treatments that would unnecessarily prolong a person's illness.

10.5.5.3 Trade-off between net health benefits and resource use

No evidence on the cost-effectiveness of interventions for adults with depression with psychotic symptoms was identified and no further economic analysis was undertaken. The GC considered the costs associated with the treatment of people with depression with psychotic symptoms, including costs of inpatient care in psychiatric wards and, potentially, of Accident and Emergency visits. The GC acknowledged that referring people with depression with psychotic symptoms to specialist mental health services was likely to incur additional costs compared with no referral, but expressed the opinion that such costs were likely to be offset by cost-savings resulting from more appropriate care for this population following referral (compared with treatment in primary care settings), leading to improved outcomes and reduction in the need for costly inpatient care. The GC assessed the costs of antipsychotics, and given that a wide range of antipsychotics are currently available in generic form, they estimated that augmentation of the current treatment plan with antipsychotic medicine was likely to lead to small resource implications.

10.5.5.4 Quality of evidence

The quality of the evidence was assessed using GRADE.

The evidence identified covered a wide range of pharmacological interventions, but was generally from single RCTs with a small sample size, and was predominantly of low to very low quality. This prevented the GC from making specific recommendations about named pharmacological interventions.

The evidence identified for combined treatment with an antidepressant and an antipsychotic when compared with various monotherapies was some of the highest quality evidence considered by the GC. This showed a greater likelihood of response and remission from illness, without unacceptable harms as evidence by side effects. They therefore agreed to retain the recommendation from the 2009 guideline to augment the current treatment plan with antipsychotic medication. Given the variable quality of the evidence and its limitations, the GC agreed that this should be a 'consider' recommendation.

Although evidence was identified relating to relapse prevention interventions in this patient group, this was much more limited than for acute treatment and came from only 4 very small RCTs of very low quality. The GC were not sufficiently confident in the findings of these studies to make any recommendations about these interventions.

10.5.5.5 Other considerations

Given the limitations of the evidence base for psychotic depression, including the fact that no evidence was identified for non-pharmacological interventions, the GC decided to develop a recommendation for further research into the most effective interventions for treatment of this condition.

10.5.6 Recommendations

94. Refer people with depression with psychotic symptoms to specialist mental health services that can provide a programme of coordinated multi-disciplinary care, which includes access to psychological interventions.[new 2017]

Update 2017

95. When treating people with depression with psychotic symptoms, consider adding antipsychotic medicine to their current treatment plan. [new 2017]

10.5.7 Research recommendations

4. What are the most effective and cost effective interventions for the treatment and management of psychotic depression (including consideration of pharmacological, psychological and psychosocial interventions)?

Statement: A series of randomised controlled trials should be conducted to determine whether pharmacological, psychological or psychosocial interventions are the most effective and cost effective at achieving remission from depression with psychotic features and improving quality of life, in adults experiencing a psychotic depressive episode.

Rationale: There is limited evidence on the most effective interventions for the treatment of psychotic depression. All identified evidence examined different pharmacological strategies, with no evidence identified for psychological or psychosocial interventions. Additionally, the current evidence for pharmacological interventions consisted primarily of small, low quality RCTs. The lack of evidence for psychological or psychosocial interventions alone or in combination with pharmacological is a further limitation. There is very little data on the long-term outcomes for people with psychotic depression. Therefore, a series of RCTs are required to compare novel pharmacological interventions and psychological and psychosocial interventions with the established treatment strategy (antidepressant treatment augmented with antipsychotic medication), to determine clinical and cost effectiveness. Follow-up should be adequate to determine the risk of relapse associated with each strategy. This study would probably require a coordinated recruitment strategy across several treatment settings and services in order to achieve adequate statistical power.

Update 2017

11 Relapse prevention

11.1 Introduction

Depression is often a recurring or chronic disorder. Although approximately half of the people who become depressed will only have a single episode of major depression in their lifetimes, approximately 50% will have multiple episodes or protracted chronic periods of depression (Eaton et al. 2008; Moffitt et al. 2010, Monroe & Harkness 2011). Among patients seeking treatment for depression, longitudinal studies find that between 50% and 85% of people with one major depressive episode will have at least one additional episode (Keller 1985). The median number of episodes reported in one large US longitudinal study was 4 (Judd et al. 1998a). Relapse is typically defined as when an individual re-experiences an episode of depression following incomplete or only brief recovery (for example less than 4 months of being well), whereas recurrence usually means a new episode following a period of recovery lasting more than 4 months, although there are only limited conceptual or evidential grounds to separate them (Frank et al. 1990).

There is robust evidence that the risk for relapse and recurrence progressively increases with each prior episode of major depression but decreases as the period of recovery is longer (Bockting et al. 2006; Solomon et al. 2000). For this reason, relapse prevention interventions, such as MBCT, have focused on individuals with a history of recurrent depression (typically defined as 2 or more lifetime episodes of major depression, but sometimes 3 or more episodes in treatment studies). Equally, individuals with a history of recurrent depression may also be more likely to relapse when withdrawn from antidepressant medication: in one study, 70% experienced a recurrence within 6 months (Frank, Kupfer and Perel 1989), raising questions about the need for continuing antidepressants beyond recovery from the acute episode.

Further predictors of relapse and recurrence include severity of initial depression, residual symptoms of depression post-initial treatment (Bockting et al. 2006; Hardeveld et al. 2010; Judd et al. 1998b; Melartin et al. 2004; Paykel et al. 1995), and a history of additional psychiatric disorder besides depression (Coryell, Endicott and Keller 1991; Melartin et al. 2004). This speaks to the potential clinical value of successfully treating residual symptoms and co-morbidity when intervening with depression, in order to maximise the likelihood of an individual staying well into the long-term. A number of variants of CBT including continuation-phase CBT, rumination-focused CBT (RFCBT) and well-being therapy have been designed to this specific goal. Since a number of randomised controlled trials of these interventions have completed since the last guideline, they will be reviewed in the context of second-line treatments and interventions for depression that has not adequately responded to treatment.

Because of the long-term nature of depression, with many patients at substantial risk of later recurrence, there is a considerable need to establish how long such patients should stay on antidepressants. The previous Guideline (NICE 2009) noted that there is strong evidence that responders to medication, who have previously had multiple relapses, should stay on medication for at least 6 months and up to 2 years after remission, to avoid relapse and recurrence, irrespective of the length of treatment pre-response (between 6 weeks and 12 months). This beneficial effect was evidenced to last beyond 12 months, but from the available data, it was not possible to determine effects beyond 2 years. A major review by Geddes and colleagues (2003) found that antidepressants reduced the risk of relapse in depression and continued treatment with antidepressants appeared to benefit many patients with recurrent depression. It was estimated that for patients who were still at appreciable risk of recurrence after 4 to 6 months of treatment with antidepressants, another year of continuation treatment would approximately halve their risk. However, there is considerable variation in practice, suggesting that many patients do not receive optimum treatment.

Update 2017

The previous Guideline (NICE 2009) noted that there is evidence that psychological treatments do not have an increased risk for relapse/recurrence following their discontinuation when compared with antidepressants, raising the possibility that some psychological interventions may confer ongoing prophylactic benefits in terms of individuals learning new coping skills and strategies that extend beyond the period of treatment. The majority of this evidence came from studies comparing CBT with antidepressants, which showed a reduced relapse rate for CBT in the follow-up of individual trials. In addition, a number of psychological interventions have been designed or adapted with a specific target of preventing relapse and recurrence including MBCT. In the light of a number of significant trials for these interventions since the last guidelines, this evidence will be reappraised in the current guideline.

11.2.2 Review question

- For adults whose depression has responded to treatment, what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions for preventing relapse (including maintenance treatment)?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 211. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 211: Clinical review protocol summary for the review of interventions for preventing relapse

Component	Description
Review question	For adults whose depression has responded to treatment, what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions for preventing relapse (including maintenance treatment)? (RQ2.3) • Does mode of delivery of psychological interventions (group-based or individual) affect outcomes? • Does format of delivery of psychological interventions (face-to-face, telephone-based or digital) affect outcomes? • Do outcomes differ for older adults?
Population	Adults whose depression has responded to treatment according to DSM, ICD or similar criteria, or depressive symptoms as indicated by depression scale score, who are randomised to relapse prevention intervention whilst in remission. Studies which explicitly targeted adults in partial remission or with residual symptoms, or the baseline depression scores did not indicate remission as decided by the GC are reviewed separately (RQ 2.4, chapter 8)
Intervention(s)	Any psychological or pharmacological intervention alone or in combination **Psychological interventions included**: • cognitive and cognitive behavioural therapies (CT / CBT) • interpersonal therapy • mindfulness-based cognitive therapy • others (CBASP) **Pharmacological interventions included:** • SSRIs • TCAs • duloxetine/venlafaxine • agomelatine

Update 2017

Component	Description
	• antipsychotics[1] • lithium augmentation **Physical interventions included:** • ECT
Comparison	• Treatment as usual • No treatment, placebo, waitlist control, attention control • Any alternative management strategy
Critical outcomes	**Psychological interventions** Relapse (the number of participants who relapsed) at: • 12 months (or closest time-point) • 24 months (or closest time-point) Sub-group analyses for psychological interventions • Remission status • Previous antidepressant use • Comparator (antidepressants or treatment as usual) • Older adults **Pharmacological interventions** • Relapse (the number of participants who relapsed) at: • Endpoint • Follow-up Sub-group analyses for pharmacological interventions • Older adults • Pre-randomisation treatment (< 6 months continuation, >6 months continuation prior to randomisation) • Randomised study time (< 12month of randomised treatment, >12 months randomised treatment) Relapse was defined a change in the status of a patient, from remitted to depressed. This had to be measured according to meeting accepted diagnostic criteria on the basis of a clinical interview, or by meeting a diagnostic threshold on a validated measure. Unvalidated or self-report measures of relapse were excluded from the analyses. ITT analysis was performed for all analyses. Discontinuation or missing data were counted as relapses.
Study design	RCTs and systematic reviews. Open label trials were not included

[1]Note that antipsychotics are not licensed for use in depression (with the exception of quetiapine which is licensed for use as an adjunctive treatment of major depressive episodes with major depressive disorder, but not as monotherapy)

Update 2017

11.3 Clinical evidence

11.3.1 Psychological interventions for relapse prevention

The Guideline Committee identified one existing systematic review (Clarke 2015) relevant to this review question which was used as a source for papers and a new search was conducted to identify papers examining psychological interventions for relapse prevention which were published after the search date of the review.

In total 24 RCTs (1 from handsearch, 22 from the Clarke 2015 systematic review, 1 from the update search) and 5 SRs were included: Bockting 2005, Bondolfi 2010, Fava 1994, Fava

1998, Frank 1990, Frank 2007, Godfrin 2010, Hollandare 2013, Jarrett 2001, Jarrett 2013, Klein 2004, Kuyken 2015, Kuyken 2008, Ma 2004, Meadows 2014, Reynolds 1999, Reynolds 2006, Segal 2010, Stangier 2013, Teasdale 2000, Wilkinson 2009, Williams 2013, Huijbers 2015, Hujbers 2016, Biesheuvel-Leliefeld 2015, Galante 2013, Gili 2015, Steinert 2014.

The included RCTs produced evidence for the following comparisons:

- psychological versus control
  - CBT/CT versus control
  - MBCT versus control
  - IPT versus control
  - 'other' psychological interventions versus control
- psychological versus psychological
  - CBT versus psychoeducation
  - IPT versus IPT
- psychological versus pharmacological
  - CBT versus antidepressants
  - IPT versus antidepressants

Eligible RCTs also contributed to the comparison of combination interventions with both psychological and pharmacological interventions, which are presented in Section 11.3.3 below.

Further information about both included and excluded studies can be found in Appendix J9. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

An intention to treat analysis was performed (which assumed dropouts as relapses).

Update 2017

11.3.1.1 Psychological interventions versus control

Eighteen RCTS compared psychological interventions against control: Bockting 2005, Bondolfi 2010, Frank 1990, Godfrin 2010, Huijbers 2015, Jarrett 2001, Jarrett 2013, Kelin 2004, Kuyken 2008, Kuyken 2015, Ma 2004, Meadows 2014, Reynolds 1999, Reynolds 2006, Segal 2010, Teasdale 2000, Wilkinson 2009 and Williams 2013.

These 18 RCTs examined four separate comparisons; CBT versus control, MBCT versus control, IPT versus control and 'other' psychological interventions versus control. In this comparison 'control' includes any non-psychological intervention as comparator; waitlist, treatment as usual or antidepressants (which would commonly be treatment as usual).

Study information can be found in Table 212 and summary of findings in Table 213, Table 214, Table 215 and Table 216.

Table 212: Study information table for trials included in the meta-analysis of psychological interventions (alone) versus control for relapse prevention for people in remission from depression

	CBT/CT vs Control	MBCT vs Control	IPT vs Control	'Other' psychological interventions vs control
Total no. of studies (N[1])	4 (557)	10 (1427)	3 (164)	1 (82)
Study ID	Bockting 2005[2]	Bondolfi 2010[6]	Frank 1990[17]	Kelin 2004

	CBT/CT vs Control	MBCT vs Control	IPT vs Control	'Other' psychological interventions vs control
	Jarrett 2001[3] Jarrett 2013[4] Wilkinson 2009[5]	Godfrin 2010[7] Huijbers 2015[8] Kuyken 2008[9] Kuyken 2015[10] Ma 2004[11] Meadows 2014[12] Segal 2010[14] Teasdale 2000[15] Williams 2013[16]	Reynolds 1999[18] Reynolds 2006[19]	
Country	Netherlands[2] US[3,4] UK[5]	Switzerland[6] Belgium[7] Netherlands[8] UK[9,10,11,16] New Zealand[12] Canada[14] Multiple[15]	USA	USA
Age (mean)	44.7 (9.5)[2] 42.7 (10.5)[3] 42.7 (11.8)[4] 74.0 (7.3)[5]	NR[6] 45.6 (10.6)[7] 51.7 (14.3)[8] 49.2 (11.2)[9] 49.5 (12.5)[10] 44.5 (8.9)[11] 48.4 (12.4)[12] 44.1 (10.9)[14] 43.3 (9.9)[15] 43.0 (12.0)[16]	40.2 (10.9)[17] NR[18] 76.8 (5.7)[19]	45.08 (11.41)
Sex (% female)	73.5%[2] 72.6%[3] 62.0%[4] NR[5]	71.6%[6] 81.1%[7] 72.0%[8,16] 76.0%[9,11,15] 77.0%[10] 81.0%[12] 63.0%[14]	75.0%[17] NR[18] 63.4%[19]	67.0%
Number of depressive episodes (mean, SD)	NR[2,5] 2.3 (0.15)[3] 3 (1.9)[4]	NR[6,7,10,15,16] 7.4 (8.8)8 6.4 (3.4)[9] 3.0 (2.0)[11] 8.1 (7.7)[12] 4.9 (2.6)[14]	NR	2.5 (1.6)
Status at randomisation	HAMD <10[2,3] HAMD≤12[4] MADRS <10[5]	MADRS≤13[6] HAMD <14[7] Full or partial remission'[8,9,10] HAMD < 10[11,14] Remission[12,15] HAMD < 7[13]	HAMD≤7[17] HAMD≤10[18,19]	HAMD-24 <=15 or decrease of >50% from baseline
Acute treatment	Antidepressants[2,5]	Antidepressants[6,8,9,10,11,12,14,15,16]	Antidepressants + IPT	CBASP, or CBASP given after non-

Update 2017

	CBT/CT vs Control	MBCT vs Control	IPT vs Control	'Other' psychological interventions vs control
	Cognitive therapy[3,4]	Unclear[7]		response to nefezadone
Intervention	Group CT + TAU: 8x weekly, 2-hour sessions in groups of 8[2] CT: 10x 1-hour sessions[3,4] Group CBT + antidepressant: 1.5 hour sessions in groups of 4 + antidepressant medication[5]	MBCT + TAU: weekly sessions of 2-2.5 hours in groups of 8-12[6,7,11,12,15,16] MBCT+AD: weekly sessions of 2-2.5 hours in groups of 8-12 plus maintenance ADM[8] MBCT+AD taper: weekly sessions of 2-2.5 hours in groups of 8-12 plus taper to placebo from antidepressants[9,10,14]	IPT + placebo: a maintenance form of IPT[17] 50 min monthly IPT sessions[18] 45 min monthly IPT sessions[19]	CBASP
Treatment length (weeks)	8 weeks[2] 35 weeks[3] 34 weeks[4] 10 weeks[5]	8 weeks[6,7,8,9,11,12,14,15,16] 52 weeks[10]	156 weeks[17] 104 weeks[18,19]	52 weeks
Comparison	TAU[2] No treatment[3] Placebo[4] Antidepressant[5]	TAU[6,7,11,12,15,16] Antidepressant[8,9,10] Taper to placebo[14]	Placebo	No treatment (assessment sessions every 4 weeks)
Follow-up length (months)	12 months[5] 24 months[2,3,4]	24 months	24 months	12 months

Notes:
N = total number of participants; TAU=treatment as usual; IPT=interpersonal therapy
[1] Number randomised, [2]Bockting 2005, [3]Jarrett 2001, [4]Jarrett 2013, [5]Wilkinson 2009, [6]Bondolfi 2010, [7]Godfrin 2010, [8]Huijbers 2015, [9]Kuyken 2008, [10]Kuyken 2015, [11]Ma 2004, [12]Meadows 2014, [14]Segal 2010, [15]Teasdale 2000, [16]Williams 2013, [17]Frank 1990, [18]Reynolds 1999, [19]Reynolds 2006

Table 213: Summary of findings for the comparison of CBT/CT versus control for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with CBT/CT (95% CI)
Relapse	471 (4 studies) 12 months	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision	RR 0.71 (0.53 to 0.95)	**Study population**	
				551 per 1000	**160 fewer per 1000** (from 28 fewer to 259 fewer)

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with CBT/CT (95% CI)
				Moderate	
				555 per 1000	**161 fewer per 1000** (from 28 fewer to 261 fewer)
Relapse	426 (3 studies) 24 months	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision	RR 0.82 (0.69 to 0.98)	**Study population**	
				713 per 1000	**128 fewer per 1000** (from 14 fewer to 221 fewer)
				Moderate	
				739 per 1000	**133 fewer per 1000** (from 15 fewer to 229 fewer)

Notes:
[1] ROB unclear or high in 1-2 domains for each study
[2] 95% ci crosses one clinical decision threshold

Table 214: Summary of findings for the comparison of MBCT versus control for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with MBCT versus control (95% CI)
Relapse	1000 (9 studies) 12 months	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision	**RR 0.79** (0.7 to 0.89)	**Study population**	
				592 per 1000	**124 fewer per 1000** (from 65 fewer to 177 fewer)
				Moderate	
				594 per 1000	**125 fewer per 1000** (from 65 fewer to 178 fewer)
Relapse	627 (2 studies) 24 months	⊕⊕⊕⊖ **moderate**[1] due to risk of bias	**RR 0.92** (0.79 to 1.08)	**Study population**	
				513 per 1000	**41 fewer per 1000** (from 108 fewer to 41 more)
				Moderate	
				535 per 1000	**43 fewer per 1000** (from 112 fewer to 43 more)

Notes:
[1] ROB unclear or high in 1-2 domains for most studies
[2] 95% CI crosses one clinical decision threshold

Table 215: Summary of findings for the comparison of IPT versus control for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with IPT (95% CI)
Relapse	193 (3 studies) 12 months	⨁⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.77** (0.63 to 0.95)	**Study population**	
				760 per 1000	**175 fewer per 1000** (from 38 fewer to 281 fewer)
				Moderate	
				783 per 1000	**180 fewer per 1000** (from 39 fewer to 290 fewer)
Relapse	187 (3 studies) 24 months	⨁⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.89** (0.74 to 1.07)	**Study population**	
				648 per 1000	**71 fewer per 1000** (from 168 fewer to 45 more)
				Moderate	
				722 per 1000	**79 fewer per 1000** (from 188 fewer to 51 more)

Notes:
[1] ROB unclear or high across multiple domains in most included studies
[2] 95% CI crosses one clinical decision threshold

Table 216: Summary of findings for the comparison of 'other' psychological interventions versus control for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Other psychological interventions (95% CI)
CBASP vs control – Relapse	82 (1 study) 12 months	⨁⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	RR 0.12 (0.02 to 0.91)	**Study population**	
				200 per 1000	**176 fewer per 1000** (from 18 fewer to 196 fewer)
				Moderate	
				200 per 1000	**176 fewer per 1000** (from 18 fewer to 196 fewer)

Notes:
[1] ROB high or unclear across multiple domains
[2] 95% CI crosses one clinical decision threshold

11.3.1.21 Psychological interventions versus psychological interventions

Two RCTS compared psychological interventions against one another: Stangier 2013 and Frank 2007.

These 2 RCTs examined four separate comparisons; CBT versus psychoeducation and IPT versus IPT (weekly versus bi-weekly, weekly versus monthly and bi-weekly versus monthly). Information on the included studies can be found in Table 217 and summary of findings in Table 218 and Table 219.

Table 217: Study information table for the comparison of psychological interventions versus psychological interventions for relapse prevention

	CBT versus psychoeducation	IPT versus IPT
Total no. of studies (N[1])	1 (180)	1 (131)
Study ID	Stangier 2013	Frank 2007
Country	Germany	USA
Age (mean)	48.6 (11.6)	30.6 (10.3)
Sex (% female)	77.2%	100%
Number of depressive episodes (mean, sd)	7.0 (7.8)	NR
Status at randomisation	HAMD<9	HAMD<=7
Acute treatment	NR	IPT plus an SSRI if needed
Intervention	CBT: 16x 50min sessions of maintenance CBT	IPT: weekly 1-hour sessions of IPT
Treatment length (weeks)	35 weeks	105 weeks
Comparison	Psychoeducation	IPT: bi-weekly or monthly 1-hour sessions
Follow-up length (months)	12 months	24 months
Note: [1]=number randomised, NR=not reported		

Table 218: Summary of findings table for the comparison of CBT versus psychoeducation for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Psychoeducation	Risk difference with CBT (95% CI)
Relapse	180 (1 study) 12 months	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision	**RR 0.85** (0.65 to 1.11)	**Study population**	
				600 per 1000	**90 fewer per 1000** (from 210 fewer to 66 more)
				Moderate	

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Psychoeducation	Risk difference with CBT (95% CI)
				600 per 1000	**90 fewer per 1000** (from 210 fewer to 66 more)

Notes:
[1] ROB unclear or high in 1-2 domains
[2] 95% CI crosses one clinical decision threshold

Table 219: Summary of findings table for the comparison of different IPT regimes for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with monthly/bi-monthly IPT	Risk difference with weekly/bi-monthly IPT (95% CI)
Relapse - Weekly IPT vs Bi-monthly IPT	87 (1 study) 24 months	⨁⨁⊖⊖ **low**[1,2] due to risk of bias, imprecision	**RR 1.24** (0.8 to 1.92)	**Study population**	
				432 per 1000	**104 more per 1000** (from 86 fewer to 397 more)
				Moderate	
				432 per 1000	**104 more per 1000** (from 86 fewer to 397 more)
Relapse - Weekly IPT vs Monthly IPT	87 (1 study) 24 months	⨁⊖⊖⊖ **very low**[1,3] due to risk of bias, imprecision	**RR 1.12** (0.74 to 1.7)	**Study population**	
				477 per 1000	**57 more per 1000** (from 124 fewer to 334 more)
				Moderate	
				477 per 1000	**57 more per 1000** (from 124 fewer to 334 more)
Relapse - Bi-monthly IPT vs monthly IPT	88 (1 study) 24 months	⨁⊖⊖⊖ **very low**[1,3] due to risk of bias, imprecision	**RR 0.9** (0.57 to 1.43)	**Study population**	
				477 per 1000	**48 fewer per 1000** (from 205 fewer to 205 more)
				Moderate	
				477 per 1000	**48 fewer per 1000** (from 205 fewer to 205 more)

Notes:
[1] ROB high or unclear across 1-2 domains

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with monthly/bi-monthly IPT	Risk difference with weekly/bi-monthly IPT (95% CI)
[2] 95% CI crosses one clinical decision threshold [3] 95% CI crosses two clinical decision thresholds					

11.3.1.31 Psychological interventions versus pharmacological interventions

2 Three RCTs (n= 288) compared psychological and pharmacological interventions: Frank
3 1990, Jarrett 2013 and Reynolds 1999.

4 These 3 RCTs examined two separate comparisons; CBT versus antidepressants and IPT
5 versus antidepressants. Information for the included studies can be found in Table 220 and
6 summary of findings in Table 221.

7 **Table 220: Study information table for trials included in the meta-analysis of**
8 **psychological interventions versus antidepressants for relapse prevention**

	CBT versus antidepressant	IPT versus antidepressant
Total no. of studies (N[1])	1 (172)	2 (116)
Study ID	Jarrett 2013	Frank 1990[2] Reynolds 1999[3]
Country	USA	USA
Age (mean)	NR	40.2 (10.9)[2] NR[3]
Sex (% female)	NR	75.0%[2] NR[3]
Number of depressive episodes (mean, sd)	NR	NR
Status at randomisation	Partial remission (HAMD>7<13)	HAMD<7[2] Remission[3]
Acute treatment	Cognitive therapy: 16 sessions	Antidepressants + IPT[2] Antidepressants[3]
Intervention	CBT: 4x 60min biweekly session then 6x monthly 60min sessions	IPT: 50min monthly sessions of IPT
Treatment length (weeks)	34 weeks	24 months
Comparison	Fluoxetine	Imipramine[2] Nortriptyline[3]
Follow-up length (months)	24 months	24 months
Notes: [1]Number randomised, [2]Frank 1990, [3]Reynolds 1999		

Table 221: Summary of findings table for trials included in the meta-analysis of psychological interventions in versus antidepressants for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Pharm	Risk difference with Psych (95% CI)
CBT vs AD - Relapse	172 (1 study) 12 months	⊕⊖⊖⊖ **very low**[1] due to risk of bias, imprecision	**RR 1** (0.73 to 1.38)	**Study population**	
				465 per 1000	**0 fewer per 1000** (from 126 fewer to 177 more)
				Moderate	
				465 per 1000	**0 fewer per 1000** (from 126 fewer to 177 more)
CBT vs AD- Relapse	155 (1 study) 24 months	⊕⊕⊖⊖ **low**[2,3] due to risk of bias, imprecision	**RR 0.88** (0.72 to 1.09)	**Study population**	
				739 per 1000	**89 fewer per 1000** (from 207 fewer to 67 more)
				Moderate	
				739 per 1000	**89 fewer per 1000** (from 207 fewer to 67 more)
IPT vs AD	115 (2 studies)	⊕⊕⊖⊖ **low**[2,3] due to risk of bias, imprecision	**RR 1.35** (0.92 to 1.98)	**Study population**	
				414 per 1000	**145 more per 1000** (from 33 fewer to 406 more)
				Moderate	
				413 per 1000	**145 more per 1000** (from 33 fewer to 405 more)

Notes:
[1] 95% CI crosses two clinical decision thresholds
[2] ROB high or unclear in 1-2 domains
[3] 95% CI crosses one clinical decision threshold

11.3.1.43 Psychological interventions for relapse prevention: subgroup analysis

At the GC's request we were able to complete subgroup analysis to examine the impact of remission status at entry, comparator treatment in the CBT/CT and MCBT analyses, as well as age of patients and acute treatment received for CBT/CT only.

The analysis showed that in individuals in partial rather than full remission at the start of the relapse prevention intervention, CBT/CT were highly effective at both 12 and 24 months for the prevention of relapses compared with control (RR 12 months=0.35 [0.20, 0.61), 24 months=0.42 [0.29, 0.62]).

Update 2017

The analysis also showed that type of control treatment (TAU, no treatment or placebo or antidepressants) had no particular impact on the efficacy of CBT/CT (RR CBT/CT versus TAU/no treatment/placebo=0.78 [0.63, 0.95], CBT/CT versus antidepressants=0.72 [0.39, 1.34]). There was also no consistent picture from the MBCT comparator sub-analysis (RR TAU=0.79 [0.69, 0.91], MBCT+AD versus AD=0.90 [0.58, 1.40], MBCT + taper versus AD=0.81 [0.62, 1.05], placebo=0.67 [0.71, 0.90]). There was no differential effect of prior treatment with either psychological therapy or an antidepressant on relapse risk with CBT/CT versus control at 12 months (RR psychological therapy=0.86 [0.65, 1.15] compared with antidepressants =0.80 [0.61, 1.04]), however at 24 months patients who had previously received antidepressants had had fewer relapses with CBT/CT as their relapse prevention intervention (RR prior ADM use=0.71 [0.57, 0.88] compared with no ADM use =0.99 [0.79, 1.22]). Finally, there did not appear to be any differential effectiveness of CBT/CT as a relapse prevention intervention in older adults versus under 65s (RR >65 years=0.53 [0.16, 1.74] compared with <65 years=0.52 [0.35, 0.76]).

11.3.2 Pharmacological interventions

In total 56 RCTs and 2 SRs were included in this review: Alexopoulos 2000, Anon 1993H, Bauer 2000, Bieling 2012, Coppen 1978a, Dobson 2008, Doogan 1992, Feiger 1999, Franchini 1998, Frank 1990, Georgotas 1989, Gilaberte 2001, Goodwin 2009, Gorwood 2007, Grunhaus 2001, Hochstrasser 2001, Hollon 2005, Jarrett 2013, Keller 1998, Kellner 2006, Kishimoto 1994, Klysner 2002, Koeser 2015, Kornstein 2006, Laurizten 1996, Lepine 2004, Liebowitz 2010, McGrath 2006, Montgomery 1988, Montgomery 1993, Montgomery 2004, Papatkos 2008, Perahia 2006, Perahia 2009, Perlis 2002, Petersen 2010, PREVENT studya, PREVENT studyb, Prien 1984, Rapaport 2004, Rapaport 2006, Reimherr 1998, Rickels 2010, Robert 1995, Robinson 1991, Rosenthal 2013, Rouillon 1991, Sackheim 2001, Schmidt 2000, Segal 2010, Shepherd 1981, Simon 2004, Stewart 1997, Terra 1998, Thase2001, van den Broek 2006, Versiani 1999, Wilson 2003

The included RCTs produced the following comparisons:

- antidepressants versus placebo
- antidepressant (full dose) versus antidepressant (half dose)
- antidepressant versus lithium
- lithium augmentation of an antidepressant versus placebo augmentation of an antidepressant
- risperidone augmentation of an antidepressant versus placebo augmentation of an antidepressant
- antipsychotics versus placebo
- responders to ECT randomised to continuation treatments.

Further information about both included and excluded studies can be found in Appendix J9. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

An intention to treat analysis was performed (which assumed dropouts as relapses).

11.3.2.1 Pharmacological interventions versus placebo

48 RCTs (n=9105) examined the relative effectiveness of antidepressant drugs compared with placebo for the prevention of relapse in depression. The drugs investigated included SSRIs (k=24; specifically sertraline, fluoxetine, fluvoxamine, escitalopram, citalopram, paroxetine and paroxetine with or without lithium or desipramine augmentation), TCAs (k=10), SNRIs (k=6) and 'other' antidepressants (k=5).

Update 2017

Information on the included studies can be found in Table 222 and Table 223, and summary of findings in Table 224.

Table 222: Study information table for trials included in the meta-analysis of antidepressants versus placebo for relapse prevention (part 1; SSRIs)

	Sertraline vs placebo	Fluoxetine vs placebo	Fluvoxamine vs placebo	Escitalopram vs placebo	Citalopram vs placebo	Paroxetine vs placebo	Paroxetine (+/- li/des) vs placebo
Total no. of studies (N[1])	6 (1081)	6 (1360)	1 (204)	3 (718)	3 (616)	3 (312)	1 (69)
Study ID	Doogan 1992[2] Jarrett 2013[3] Keller 1998[4] Lepine 20045 Segal 20106 Wilson 20037	Gilaberte 20018McGrath 2006[10] Montgomery 1988[11] Petersen 2010[12] Reimherr 1998[13] Schmidt 2000[14]	Terra 1998	Gorwood 2007[15] Kornstein 2006[16] Rapaport 2004[17]	Hochstrasser 2001[18] Klysner 2002[19] Robert 1995[21]	Dobson 2008[22] Montgomery 1993[20] Reynolds 2006[23]	Hollon 2005
Country	NR[2,4,7] USA[3] France[5] Canada[6]	NR[8,14]Europe and USA[10,11] USA[12]	NR	Czech Republic, France, Germany, Netherlands, Poland, Slovakia, Spain[15] USA[16,17]	NR	Canada[22] NR[20] USA[23]	USA
Age (mean)	NR[2,3,4] 48.0 (11.2)[5] 44.0 (11.0)[6] 77.7[7]	NR[8,11,14]38.0[10] 39.9 (10.3)[12]	NR	73.0[15] 43.0[16] 42.0[17]	NR	38.9 (10.0)[22] NR[20] 76.8 (5.7)[23]	NR
Sex (% female)	NR[2,3,4,7] 67.8%[5] 65.3%[6]	NR[8,11,14]54.9%[10] 55.0%[12]	NR	78.7%[15] 79.1%[16] 60.9%[17]	NR	78.2%[22] NR[20] 63.8%[23]	NR
Acute treatment	Sertraline[2,7] Cognitive therapy[3] Sertraline or imipramine[4] NR[5]	Fluoxetine	Fluvoxamine	Escitalopram[15,17] SSRI[16]	Citalopram	Paroxetine[22,20] Paroxetine + weekly IPT[23]	Antidepressant

Update 2017

	Sertraline vs placebo	Fluoxetine vs placebo	Fluvoxamine vs placebo	Escitalopram vs placebo	Citalopram vs placebo	Paroxetine vs placebo	Paroxetine (+/- li/des) vs placebo
	Pharmacotherapy algorithm[6]						
Treatment length	10 months[2] 8 months[3] 18 months[4,5,6] 24 months[7]	25 weeks[14] 11 months[8]12 months[10,11,13] 80 weeks[12]	12 months	6 months[15] 12 months[16] 8 months[17]	48 weeks[18,19] 24 weeks[21]	12 months[20,22]	12 months
Intervention	Sertraline: 50-200mg/day[2]; dose NR[3,4,6]; 100mg/day[5]; 50-100mg/day[7]	Fluoxetine: 20mg/day[8,12,14] dose NR[10,13] 40mg/day[11]	Fluvoxamine 100mg/day	Escitalopram: 10-20mg/day[15,17]; dose NR[16]	Citalopram: 20-60mg/day[18,21]; 20-40mg/day[19]	Paroxetine: continuation dose[22]; 20-30mg/day[20]; 10-40mg/day with clinical management[23]	Paroxetine (+/-) lithium or desipramine augmentation
Comparison	Placebo	Placebo	Placebo	Placebo	Placebo	Placebo	Placebo

Notes:

[1] Number randomised. NR=not reported

[2]Doogan 1992, [3]Jarrett 2013, [4]Keller 1998, [5]Lepine 2004, [6]Segal 2010, [7]Wilson 2003, [8]Gilaberte 2001, [10]McGrath 2006, [11]Montgomery 1988, [12]Petersen 2010, [13]Reimherr 1998, [14]Schmidt 2000, [15]Gorwood 2007, [16]Kornstein 2006, [17]Rapaport 2004, [18]Hochstrasser 2001, [19]Klysner 2002, [20]Montgomery 1993, [21]Robert 1995, [22]Dobson 2008, [23]Reynolds 2006

Update 2017

Table 223: Study information table for trials included in the meta-analysis of antidepressants versus placebo for relapse prevention (part 2; TCAs, SNRIs, 'other' antidepressants)

	TCAs vs Placebo	SNRIs vs Placebo	'Other' antidepressants vs Placebo
Total no. of studies (N[1])	10 (1,246)	6 (1,450)	5 (829)
Study ID	Alexopolous 2000[2] Anon 1993H[3] Coppen 1978a[4] Frank 1990[5] Georgotas 1989[6] Prien 1984[7]	Montgomery 2004[13] Perahia 2006[14] Perahia 2009[15] PREVENT study a[16] PREVENT study b[17] Simon 2004[18]	Feiger 1999[19] Goodwin 2009[20] Kishimoto 1994[21] Robinson 1991[22] Versiani 1999[23]

	TCAs vs Placebo	SNRIs vs Placebo	'Other' antidepressants vs Placebo
	Rouillon 1991[9] Sackeim 2001[10] Stewart 1997[11] van den Broek 2006[12]		
Country	UK[2,3] NR[4,5,6,7] France[9] USA[10,11] Ntherlands[12]	Europe and USA[13] Italy, France, Spain, US[14] France, Germany, Italy, Russia, Sweden[15] USA[16,17,18]	NR[19,21,22,23] Australia, Finland, France, South Africa, UK[20]
Age (mean)	65.0[2] 75.7 (6.2)[3] 55.3[4] NR[5,7] 64.0[6] 46.0 (12)[9] 57.0[10] 39.0 (8)[11] 51.0[12]	43.8 (11)[13] 45.0[14] 47.1 (12.8)[15] 42.0[16,17] 43.0[18]	NR[19,21,22,23] 43.4 (10.9)[20]
Sex (% female)	73.0%[3] 81.3%[4] NR[5,7] 70.0%[9] 66.6%[10] 57.0%[11] 74.1%[12]	71.0%[13] 72.7%[14] 68.5%[15] 68.2[16,17] 59.1%[18]	NR[19,21,22,23] 72.1%[20]
Acute treatment	Nortriptyline[2] Any considered suitable[3,4,7] Imipramine + IPT[5] Phenelzine or imipramine[6] Maprotiline[9] ECT[10] Phenelzine or imipramine[11]	Venlafaxine IR[13,16,17] Duloxetine[14,15] Venlafaxine XR 150 or 225 mg[18]	Nefazodone[19] Agomelatine 25-50mg/day[20] TCA or mianserin[21] Phenelzine[22] Reboxetine[23]

Update 2017

	TCAs vs Placebo	SNRIs vs Placebo	'Other' antidepressants vs Placebo
Treatment length	24 months[2,3,7] 12 months[4,6,9] 3 years[5] 6 months[10,11,12]	12 months[13,15,16,17] 6 months[14,18]	36 weeks[19] 24 weeks[20] 18 months[21] 24 months[22] 46 weeks[23]
Intervention	Nortriptyline[2,6,10] Dothiepin: 75mg/day[3] Amitriptyline[4] Maprotiline: 75mg/day[9] Imipramine: continuation dose[5,7,11,12]	Venlafaxine: 100-200mg/day[13]; ER dose NR[16,17]; 72-225mg/day[18] Duloxetine: dose NR[14]; 60-120mg/day[15]	Nefazodone[19] Agomelatine 25-50mg/day[20] Mianserin 24-26mg/day[21] Phenelzine 45-60mg/day[22] Reboxetine 8mg/day[23]
Comparison	Placebo	Placebo	Placebo

Notes:

[1] Number randomised

[2]Alexopolous 2000, [3]Anon 1993H, [4]Coppen 1978a, [5]Frank 1990, [6]Georgotas 1989, [7]Prien 1984, [9]Rouillon 1991, [10]Sackeim 2001, [11]Stewart 1997, [12]van den Broek 2006, [13]Montgomery 2004, [14]Perahia 2006, [15]Perahia 2009, [16]PREVENT study a, [17]PREVENT study b, [18]Simon 2004, [19]Feiger 1999, [20]Goodwin 2009, [21]Kishimoto 1994, [22]Robinson 1991, [23]Versiani 1999

Note that maprotiline and nefazodone are not available in the UK but have been included in this review in order to assess the class effect of pharmacological interventions for depression

Update 2017

1

Table 224: Summary of findings for the comparison of antidepressant versus placebo for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Antidepressant (95% CI)
Relapse-All	9105 (48 studies)	⊕⊖⊖⊖ **VERY LOW**[1,2] due to risk of bias, inconsistency	**RR 0.59** (0.55 to 0.65)	**Study population**	
				524 per 1000	**215 fewer per 1000** (from 184 fewer to 236 fewer)
				Moderate	
				524 per 1000	**215 fewer per 1000** (from 183 fewer to 236 fewer)

[1] ROB low in only one or two domains and downgraded twice accordingly
[2] I2 >50% <80%

Subgroup analysis: antidepressants versus placebo

Subgroup analysis was possible to investigate differential effects of different drugs, the impact of age and of post-randomisation time on the effectiveness of antidepressants as a relapse prevention intervention.

The effect of antidepressants relative to placebo on relapse prevention held across all drug classes and almost all drugs within the SSRI class. In SSRIs the relative risks were 0.57 (0.49, 0.67) for the class overall, 0.65 (0.46, 0.93) for sertraline, 0.56 (0.49, 0.64) for fluoxetine, 0.36 (0.21, 0.64) for fluvoxamine, 0.44, (0.25, 0.75) for escitalopram, 0.47 (0.36, 0.61) for citalopram, 0.80 (0.59, 1.07) for paroxetine and 0.61 (0.41, 0.91) for paroxetine with or without lithium or desipramine augmentation. For TCAs the relative risk was 0.59 (0.46, 0.75), for SNRIs the RR was 0.55 (0.44, 0.69) and for 'other antidepressants' the RR was 0.41 (0.33, 0.52).

In older adults the effectiveness of antidepressants compared to placebo was greater, but less certain, than in the overall comparison (RR older adults=0.46 [0.30, 0.72] versus overall=0.59 [0.56, 0.64]).

Duration of follow-up (time since randomisation) did not appear to have a significant impact upon the findings of this comparison with relative risks of 0.52 (0.43, 0.65) in studies with under 12 months of follow-up, 0.56 (0.48, 0.65) in studies with between 12 and 24 month follow-up and 0.57 (0.41, 0.78) in studies with over 24 month follow-up.

11.3.2.2 Full dose antidepressant treatment versus half dose antidepressant treatment for the prevention of relapse

3 RCTs (n=1024, Franchini 1998, Lepine 2004 and Rouillon 1991) examined the relative effectiveness of receiving a full versus half dose of antidepressant drugs for the prevention of relapse in depression. The drugs investigated included SSRIs and TCAs.

Information on the included studies can be found in Table 225, and summary of findings in Table 226.

Update 2017

Table 225: Study information table for trials included in the meta-analysis of full versus half dose of antidepressants for relapse prevention

	Antidepressant (full dose) versus antidepressant (half dose)
Total no. of studies (N[1])	3 (1,024)
Study ID	Franchini 19982 Lepine 20043 Rouillon 19914
Country	Italy2 France3,4
Age (mean)	47.0 (8.9)[2] NR[3] 46.0 (12.0)[4]
Sex (% female)	64.7%[2] 67.8%[3] 70.0%[4]
Acute treatment	Paroxetine 40mg/day[2] Antidepressant[3] Maprotiline[4]
Treatment length	28 months[2] 18 months[3] 12 months[4]
Intervention	Paroxetine 40mg/day[2] Sertraline 100mg/day[3] Maprotiline 75mg/day[4]
Comparison	Paroxetine 20mg/day[2] Sertraline 50mg/day[3] Maprotiline 37.5mg/day
Notes: [1]Number randomised [2]Franchini 1998, [2]Lepine 2004, [4]Rouillon 1991 Note that maprotiline is not available in the UK but has been included in this review to assess the class effect of pharmacological interventions for depression	

Table 226: Summary of findings for the comparison of antidepressant (full dose) versus antidepressant (half dose) for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Antidepressant (half dose)	Risk difference with Antidepressant (full dose) (95% CI)
Relapse (any antidepressant: full versus half dose)	1024 (3 studies)	⊕⊖⊖⊖ **very low**[1,2,3] due to risk of bias, inconsistency, imprecision	**RR 0.81** (0.60 to 1.08)	**372 per 1000**	**71 fewer per 1000** (from 149 fewer to 1000 more)
Relapse: TCA (full dose) versus TCA (half dose)	767 (1 study)	⊕⊕⊖⊖ **low**[3,4] due to risk of bias, imprecision	**RR 0.80** (0.65 to 0.99)	**356 per 1000**	**71 fewer per 1000** (from 4 fewer to 125 fewer)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Antidepressant (half dose)	Risk difference with Antidepressant (full dose) (95% CI)
Relapse: SSRI (full dose) versus SSRI (half dose)	257 (2 studies)	⨁⊖⊖⊖ **very low**[1,3] due to risk of bias, imprecision	**RR 0.73** (0.35 to 1.04)	**419 per 1000**	**113 fewer per 1000** (from 272 fewer to 226 more)

Notes:
[1] ROB high or unclear across multiple domains
[2] I-squared >50%<80%
[3] 95% CI crosses one clinical decision threshold
[4] ROB unclear in several domains

11.3.2.31 Antidepressants versus lithium for relapse prevention

1 RCT (n=107; Shepherd 1981) examined the relative effectiveness of antidepressants versus lithium for the prevention of relapse in depression. Information on the included studies can be found in Table 227 and summary of findings in Table 228.

Table 227: Study information table for the comparison of antidepressants versus lithium for relapse prevention

	Antidepressants versus lithium
Total no. of studies (N[1])	1 (107)
Study ID	Shepherd 1981
Country	UK
Age (mean)	NR
Sex (% female)	81.0%
Status at randomisation	Remission
Acute treatment	NR
Treatment length	3 years
Intervention	Amitriptyline (dose NR)
Comparison	Lithium (dose NR)

Notes:
[1]Number randomised, NR=not reported

Table 228: Summary of findings for the comparison of antidepressants versus lithium for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Lithium alone	Risk difference with Antidepressant (95% CI)
Relapse - Amitriptyline vs lithium	107 (1 study)	⨁⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.72** (0.55 to 0.95)	**Study population**	
				780 per 1000	**218 fewer per 1000** (from 39 fewer to 351 fewer)
				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Lithium alone	Risk difference with Antidepressant (95% CI)
				780 per 1000	**218 fewer per 1000** (from 39 fewer to 351 fewer)

Notes:
[1] ROB high or unclear across multiple domains
[2] 95% CI crosses one clinical decision threshold

11.3.2.41 Lithium augmentation of antidepressants versus placebo for relapse prevention

2 3 RCTs (n=160; Bauer 2000, Prien 1984, Sackeim 2001) examined the relative effectiveness
3 of lithium augmentation of antidepressants versus placebo augmentation of antidepressants
4 for the prevention of relapse in depression.

5 Information on the included studies can be found in Table 229 and summary of findings in
6 Table 230.

7 **Table 229: Study information table for trials included in the meta-analysis of lithium**
8 **augmentation of antidepressants versus placebo augmentation of**
9 **antidepressants for relapse prevention**

	Lithium augmentation versus placebo augmentation
Total no. of studies (N[1])	3 (160)
Study ID	Bauer 2000[2] Prien 1984[3] Sackeim 2001[4]
Country	NR2,3 USA4
Age (mean)	47.42 NR3 57.04
Sex (% female)	NR[2,3] 66.7%[4]
Acute treatment	Antidepressant (+ lithium augmentation if no response after 4 weeks)[2] Any clinician deemed appropriate[3] ECT[4]
Treatment length	4 months[2] 24 months[3] 6 months[6]
Intervention	Antidepressant + lithium augmentation; doses NR[2,3] Nortriptyline + lithium[4]
Comparison	Antidepressant + placebo; doses NR[2,3] Nortriptyline + placebo[4]

Notes:
[1]Number randomised
[2]Bauer 2000, [3]Prien 1984, [4]Sackeim 2001

Table 230: Summary of findings for lithium augmentation of antidepressants versus placebo augmentation of antidepressants

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo + AD	Risk difference with Lithium augmentation + AD (95% CI)
Relapse	160 (3 studies)	⨁⨁⊖⊖ **low**[1,2] due to risk of bias, imprecision	RR 0.62 (0.35 to 1.12)	**Study population**	
				412 per 1000	**157 fewer per 1000** (from 268 fewer to 49 more)
				Moderate	
				385 per 1000	**146 fewer per 1000** (from 250 fewer to 46 more)

Notes:
[1] ROB high or unclear in several domains
[2] 95% CI crosses one clinical decision threshold

11.3.2.5 Risperidone augmentation of antidepressants versus placebo augmentation of antidepressants for relapse prevention

1 RCT (n=241; Rapaport 2006) examined the relative effectiveness of risperidone augmentation of antidepressants versus placebo augmentation of antidepressants for the prevention of relapse in depression.

Information on the included studies can be found in Table 231 and summary of findings in Table 232.

Table 231: Study information table for the comparison of risperidone augmentation of antidepressants versus placebo augmentation of antidepressants for relapse prevention

	Risperidone augmentation versus placebo augmentation
Total no. of studies (N[1])	1 (241)
Study ID	Rapaport 2006
Country	USA
Age (mean)	48.0
Sex (% female)	60.9%
Acute treatment	Citalopram
Treatment length	6 months
Intervention	Citalopram + risperidone
Comparison	Citalopram + placebo
Note: [1]Number randomised	

Table 232: Summary of findings for risperidone augmentation of antidepressants versus placebo augmentation of antidepressants

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo + AD	Risk difference with Risperidone augmentation + AD (95% CI)
Relapse	241 (1 study)	⨁⨁⊖⊖ **LOW**[1,2] due to risk of bias, imprecision	**RR 0.98** (0.77 to 1.23)	**Study population**	
				546 per 1000	**11 fewer per 1000** (from 126 fewer to 126 more)
				Moderate	
				546 per 1000	**11 fewer per 1000** (from 126 fewer to 126 more)

Notes:
[1] ROB unclear across several domains
[2] OIS not met (<300 events)

11.3.2.63 Antipsychotics versus placebo for relapse prevention

1 RCT (n=776; Liebowitz 2010) examined the effectiveness of antipsychotics relative to placebo for the prevention of relapse in depression.

Information on the included studies can be found in Table 233 and summary of findings in Table 234.

Table 233: Study characteristics for trials included in the meta-analysis of antipsychotics versus placebo for relapse prevention

	Antipsychotics versus placebo
Total no. of studies (N[1])	1 (776)
Study ID	Liebowitz 2010
Country	Bulgaria, Finland, France, Germany, Romania, Russia, the Slovak Republic, UK, Canada, South Africa, USA
Age (mean)	43.8 (11.5)
Sex (% female)	66.1%
Status at randomisation	MADRS<=12 + CGI<=3
Acute treatment	Quetiapine XR: 50mg days 1-2, increasing to 150mg on days 3-4, increasing on day 5 to 300mg if necessary
Treatment length (weeks)	52 weeks
Intervention	Quetiapine: 50-300mg/day
Comparison	Placebo
Note: [1]Number randomised	

Update 2017

Table 234: Summary of findings for the comparison of antipsychotics versus placebo for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antipsychotics versus placebo (95% CI)
Relapse - Quetiapine versus placebo	771 (1 study)	⊕⊕⊖⊖ **low**[1] due to risk of bias	**RR 0.42** (0.32 to 0.56)	**Study population**	
				331 per 1000	**192 fewer per 1000** (from 146 fewer to 225 fewer)
				Moderate	
				331 per 1000	**192 fewer per 1000** (from 146 fewer to 225 fewer)

Notes:
[1] ROB high or unclear across multiple domains

11.3.2.7 Relapse prevention interventions for individuals who responded to acute treatment with ECT

5 RCTs (n=515; Grunhaus 2001, Kellner 2006, Lauritzen 1996, Nordenskjold 2012, Sackeim 2001) examined the effectiveness of various strategies for the prevention of relapse in individuals who had responded to acute treatment with ECT for their depression.

Information on the included studies can be found in Table 235 and summary of findings in Table 236.

Table 235: Study information table for the comparison of relapse prevention interventions in ECT responders

	Antidepressant (+/- lithium) versus Placebo	Antidepressant versus Antidepressant	Antidepressant versus Antidepressant + melatonin	ECT (+/- antidepressant) versus Antidepressant (+/- lithium)
Total no. of studies (N[1])	2 (84)	2 (98)	1 (39)	2 (257)
Study ID	Lauritzen 1996[2] Sackeim 2001[3]	Lauritzen 1996[2] Sackeim 2001[3]	Grunhaus 2001	Kellner 2006[4] Nordenskjold 2012[5]
Country	Denmark[2] USA[3]	Denmark[2] USA[3]	Israel	Sweden[5]
Age (mean)	59.0[2] 57.0[3]	59.0[2] 57.0[3]	60.0	NR[5]
Sex (% female)	74.3%[2] 66.7%[3]	74.3%[2] 66.7%[3]	62.9%	50%[5]
Status at randomisation	NR[2] HAMD-24 <10 or 60% reduction from baseline[3]	NR[2] HAMD-24 <10 or 60% reduction from baseline[3]	HAMD<=10	MADRS<=10[5]
Acute treatment	ECT	ECT	ECT	ECT

Update 2017

	Antidepressant (+/- lithium) versus Placebo	Antidepressant versus Antidepressant	Antidepressant versus Antidepressant + melatonin	ECT (+/- antidepressant) versus Antidepressant (+/- lithium)
Treatment length	20 weeks[2] 6 months[3]	20 weeks[2] 6 months[3]	12 weeks	52 weeks[5]
Intervention	Imipramine 20-60mg/day[2] Nortriptyline + lithium[3]	Imipramine 20-60mg/day[2] Nortriptyline + lithium[3]	Fluoxetine + placebo: 20-40mg/day fluoxetine plus placebo	ECT + pharmacotherapy: 29x ultrabrief pulse ECT sessions plus individualised pharmacotherapy[5]
Comparison	Placebo	Paroxetine 20-60mg/day[2] Nortriptyline[3]	Fluoxetine + melatonin: 20-40mg/day fluoxetine plus 5-10mg/day melatonin	Pharmacotherapy[5]
Notes: [1]Number randomised [2]Lauritzen 1996, [3]Sackeim 2001, [4]Kellner 2006, [5]Nordenskjold 2012				

Table 236: Summary of findings for the comparison of relapse prevention interventions in ECT responders

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Responders to ECT randomised to continuation treatments (95% CI)
Relapse - ECT vs nortriptyline + lithium (after 6 months)	201 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	RR 1.16 (0.77 to 1.74)	**Study population**	
				291 per 1000	**47 more per 1000** (from 67 fewer to 216 more)
				Moderate	
				291 per 1000	**47 more per 1000** (from 67 fewer to 215 more)
Relapse - Fluoxetine + placebo vs fluoxetine + melatonin (after 12 weeks)	39 (1 study)	⊕⊖⊖⊖ **very low**[1,3] due to risk of bias, imprecision	RR 1.17 (0.4 to 3.39)	**Study population**	
				238 per 1000	**40 more per 1000** (from 143 fewer to 569 more)
				Moderate	
				238 per 1000	**40 more per 1000** (from 143 fewer to 569 more)
				Study population	

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Responders to ECT randomised to continuation treatments (95% CI)
Relapse - Nortriptyline + lithium vs placebo (after 6 months)	57 (1 study)	⊕⊕⊖⊖ **low**[2,4] due to risk of bias, imprecision	**RR 0.44** (0.25 to 0.8)	**724 per 1000**	**406 fewer per 1000** (from 145 fewer to 543 fewer)
				Moderate	
				724 per 1000	**405 fewer per 1000** (from 145 fewer to 543 fewer)
Relapse - Nortriptyline vs placebo (after 6 months)	56 (1 study)	⊕⊕⊖⊖ **low**[2,4] due to risk of bias, imprecision	**RR 0.77** (0.51 to 1.15)	**Study population**	
				724 per 1000	**167 fewer per 1000** (from 355 fewer to 109 more)
				Moderate	
				724 per 1000	**167 fewer per 1000** (from 355 fewer to 109 more)
Relapse - Nortriptyline + lithium vs nortriptyline (after 6 months)	56 (1 study)	⊕⊕⊖⊖ **low**[2] due to risk of bias, imprecision	**RR 0.6** (0.32 to 1.14)	**Study population**	
				536 per 1000	**214 fewer per 1000** (from 364 fewer to 75 more)
				Moderate	
				536 per 1000	**214 fewer per 1000** (from 364 fewer to 75 more)
Relapse - Imipramine vs paroxetine (after 6 months)	43 (1 study)	⊕⊖⊖⊖ **very low**[2,5] due to risk of bias, imprecision	**RR 3.82** (0.91 to 15.95)	**Study population**	
				95 per 1000	**269 more per 1000** (from 9 fewer to 1000 more)
				Moderate	
				95 per 1000	**268 more per 1000** (from 9 fewer to 1000 more)
Relapse - Imipramine vs placebo (after 6 months)	27 (1 study)	⊕⊖⊖⊖ **very low**[2,5] due to risk of bias, imprecision	**RR 0.21** (0.06 to 0.76)	**Study population**	
				800 per 1000	**632 fewer per 1000** (from 192 fewer to 752 fewer)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Responders to ECT randomised to continuation treatments (95% CI)
				Moderate	
				800 per 1000	**632 fewer per 1000** (from 192 fewer to 752 fewer)
Relapse - ECT + ADM vs ADM alone (12 months)	56 (1 study)	⊕⊕⊖⊖ **low**[2,4] due to risk of bias, imprecision	**RR 0.53** (0.29 to 0.98)	**Study population**	
				607 per 1000	**285 fewer per 1000** (from 12 fewer to 431 fewer)
				Moderate	
				607 per 1000	**285 fewer per 1000** (from 12 fewer to 431 fewer)

Notes:
[1] ROB high or unclear across multiple domains
[2] 95% CI crosses one clinical decision threshold
[3] 95% CI crosses two clinical decision thresholds
[4] ROB unclear across several domains
[5] ROB low in only two domains

11.3.3 Combination interventions for relapse prevention

RCTs for this review came from both the psychological and pharmacological reviews. Eight RCTs provided data for combination treatments compared either with a pharmacological monotherapy or with a psychological monotherapy: Frank 1990, Huijbers 2015, Huijbers 2016, Perlis 2002, Petersen 2010, Reynolds 1999, Reynolds 2006 and Wilkinson 2009.

11.3.3.1 Combination psychological and pharmacological interventions versus pharmacological interventions

5 RCTs (n=361) provided data for this review: Frank 1990, Huijbers 2015, Perlis 2002, Reynolds 2006 and Wilkinson 2009.

These RCTs produced three different comparisons; CBT in combination with antidepressants versus antidepressants alone, IPT in combination with antidepressants versus antidepressants alone, and MBCT in combination with antidepressants versus antidepressants alone.

Information on the included studies can be found in Table 237 and summary of findings in Table 238.

Table 237: Study information table for trials included in the meta-analysis of psychological interventions in combination with antidepressants versus antidepressants for relapse prevention

	CBT + AD vs AD	MBCT + AD versus AD	IPT+ antidepressant versus Antidepressant
Total no. of studies (N[1])	2 (177)	1 (68)	2 (116)
Study ID	Perlis 2002[2] Wilkinson 2009[3]	Huijbers 2015	Frank 1990[4] Reynolds 2006[5]
Country	USA[2] UK[3]	Netherlands	USA
Age (mean)	38.8 (10.6)[2] 74.0 (7.3)[3]	51.7 (14.3)[8]	40.2 (10.9)[4] 76.8 (5.7)[5]
Sex (% female)	58.0%[2] NR[3]	72.0%	75.0%[4] 63.4%[5]
Number of depressive episodes (mean, sd)	5.6 (9.2)[2] NR[3]	7.4 (8.8)	NR
Status at randomisation	Remission[2] MADRS <10[3]	Full or partial remission	HAMD≤7[4] HAMD≤10[5]
Acute treatment	Fluoxetine 20mg/day[2] Antidepressants[3]	Antidepressants	Antidepressants + IPT
Treatment length (weeks)	28 weeks[2] 10 weeks[3]	8 weeks	156 weeks[4] 104 weeks[5]
Intervention	CT + fluoxetine: dose increased from 20mg/day to 40mg/day; CT 12x weekly then 7xbi-weekly sessions[2] Group CBT + antidepressant: 1.5 hour sessions in groups of 4 + antidepressant medication[3]	MBCT+AD: weekly sessions of 2-2.5 hours in groups of 8-12 plus maintenance ADM	IPT + imipramine: monthly IPT sessions + imipramine 200mg/day[4] IPT + paroxetine: 45min monthly IPT sessions plus 10-40mg/day paroxetine[5]
Comparison	Fluoxetine increased from 20-40mg/day plus medication management[2] Antidepressant[3]	Antidepressant	Imipramine 200mg/day[4] Paroxetine 10-40mg/day[5]
Follow-up length	12 months	24 months	24 months
Notes: [1]Number randomised, [2]Perlis 2002, [3]Wilkinson 2009, [4]Frank 1990, [5]Reynolds 2006			

Update 2017

Table 238: Summary of findings table for the comparison of combination psychological and pharmacological interventions versus pharmacological interventions

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Pharm- 12 month	Risk difference with Combination pharm + Psych (95% CI)
Relapse **Imipramine + IPT vs Imipramine**	53 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.41** (0.15 to 1.12)	**Study population**	
				393 per 1000	**232 fewer per 1000** (from 334 fewer to 47 more)
				Moderate	
				393 per 1000	**232 fewer per 1000** (from 334 fewer to 47 more)
Relapse **MBCT + AD vs AD**	68 (1 study)	⊕⊖⊖⊖ **very low**[1,3] due to risk of bias, imprecision	**RR 0.9** (0.58 to 1.4)	**Study population**	
				571 per 1000	**57 fewer per 1000** (from 240 fewer to 229 more)
				Moderate	
				571 per 1000	**57 fewer per 1000** (from 240 fewer to 228 more)
Relapse **Paroxetine + IPT vs paroxetine**	63 (1 study)	⊕⊖⊖⊖ **very low**[3,4] due to risk of bias, imprecision	**OR 0.86** (0.42 to 1.4)	**Study population**	
				457 per 1000	**37 fewer per 1000** (from 196 fewer to 84 more)
				Moderate	
				457 per 1000	**37 fewer per 1000** (from 196 fewer to 84 more)
Relapse **CBT vs AD vs AD alone**	177 (2 studies)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	**RR 0.86** (0.62 to 1.21)	**472 per 1000**	**66 fewer per 1000** (from 179 fewer to 99 more)
Relapse **CT + fluoxetine versus fluoxetine alone**	132 (1 study)	⊕⊖⊖⊖ **very low**[1,3] due to risk of bias, imprecision	**RR 0.93** (0.62 to 1.39)	**439 per 1000**	**31 fewer per 1000** (from 167 fewer to 171 more)
Relapse				**Study population**	

Update 2017

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Pharm- 12 month	Risk difference with Combination pharm + Psych (95% CI)
CBT + AD vs AD	45 (1 study)	⊕⊖⊖⊖ **very low**[3,4] due to risk of bias, imprecision	**RR 0.72** (0.3 to 1.23)	**565 per 1000**	**158 fewer per 1000** (from 396 fewer to 130 more)
				Moderate	
				565 per 1000	**158 fewer per 1000** (from 396 fewer to 130 more)

[1] ROB high or unclear across multiple domains
[2] 95% CI crosses one clinical decision threshold
[3] 95% CI crosses two clinical decision thresholds
[4] ROB high or unclear across 1-2 domains

11.3.3.21 Combination psychological and pharmacological versus psychological interventions

5 RCTs (n=447) provided data for this review: Frank 1990, Huijbers 2016, Petersen 2010, Reynolds 1999 and Reynolds 2006.

These RCTs produced three different comparisons; CBT in combination with fluoxetine versus CBT alone, IPT in combination with antidepressants versus IPT alone, and MBCT in combination with antidepressants versus MBCT alone.

Information on the included studies can be found in Table 239 and summary of findings in Table 240.

Table 239: Study information table for the comparison of combined psychological and pharmacological interventions versus psychological interventions alone for relapse prevention

	CBT + fluoxetine versus CBT	IPT + antidepressant versus IPT	MBCT + antidepressant versus MBCT
Total no. of studies (N[1])	1 (22)	3 (176)	1 (249)
Study ID	Petersen 2010	Frank 1990[2] Reynolds 1999[3] Reynolds 2006[4]	Huijbers 2016
Country	USA	USA	Netherlands
Age (mean)	39.9 (10.3)	40.2 (10.9)[2] NR[3] 76.8 (5.7)[4]	50.3 (10.6)
Sex (% female)	55.0%	75.0%[2] NR[3] 63.4%[4]	67.0%
Number of depressive episodes (mean, SD)	5.4 (4.5)	NR	5.9 (5.3)
Status at randomisation	HAMD<=7	HAMD≤7[2] HAMD≤10[3,4]	Partial remission

Update 2017

	CBT + fluoxetine versus CBT	IPT + antidepressant versus IPT	MBCT + antidepressant versus MBCT
Acute treatment	Fluoxetine 20mg/day for 8 weeks	Antidepressants + IPT	Antidepressant treatment for at least 6 months
Treatment length (weeks)	20 months	156 weeks[2] 104 weeks[3,4]	8 weeks
Intervention	CBT + fluoxetine: 7x biweekly, 50-min sessions followed by 16x monthly, 50-min sessions plus maintenance fluoxetine	IPT + imipramine: maintenance IPT plus 200mg/day imipramine[2] IPT + nortriptyline: monthly 50min sessions plus imipramine[3] IPT + paroxetine: monthly 45 min sessions plus 10-40mg/day paroxetine[4]	MBCT + antidepressant: 8x weekly 2.5 hour sessions in groups of 8-12 plus antidepressant medication
Comparison	CBT + placebo: 7x biweekly, 50-min sessions followed by 16x monthly, 50-min sessions plus placebo pills	IPT: maintenance IPT[2] monthly 50min sessions[3] 45 min sessions[4]	MBCT: 8x weekly 2.5 hour session, with medication withdrawn over 5 weeks
Follow-up length	20 months	24 months	12 months
Notes: [1]Number randomised [2]Frank 1990, [3]Reynolds 1999, [4]Reynolds 2006			

Table 240: Summary of findings for the comparison of combined psychological and pharmacological interventions versus psychological interventions alone for relapse prevention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Psych-12 month	Risk difference with Combination pharm + psych (95% CI)
CBT + fluoxetine vs CBT	24 (1 study)	⨁⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	RR 0.79 (0.3 to 2.09)	**Study population**	
				462 per 1000	**97 fewer per 1000** (from 323 fewer to 503 more)
				Moderate	
				462 per 1000	**97 fewer per 1000** (from 323 fewer to 504 more)
				Study population	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Psych-12 month	Risk difference with Combination pharm + psych (95% CI)
IPT + Imipramine vs IPT	51 (1 study)	⊕⊖⊖⊖ **very low**[3,4] due to risk of bias, imprecision	RR 0.3 (0.11 to 0.78)	**538 per 1000**	**377 fewer per 1000** (from 118 fewer to 479 fewer)
				Moderate	
				539 per 1000	**377 fewer per 1000** (from 119 fewer to 480 fewer)
IPT + nortriptyline vs IPT	60 (1 study)	⊕⊕⊖⊖ **low**[1,4] due to risk of bias, imprecision	**RR 0.65** (0.38 to 1.14)	**Study population**	
				581 per 1000	**203 fewer per 1000** (from 360 fewer to 81 more)
				Moderate	
				581 per 1000	**203 fewer per 1000** (from 360 fewer to 81 more)
IPT + paroxetine vs IPT	63 (1 study)	⊕⊕⊖⊖ **low**[1,4] due to risk of bias, imprecision	**RR 0.6** (0.36 to 1)	**Study population**	
				657 per 1000	**263 fewer per 1000** (from 421 fewer to 0 more)
				Moderate	
				657 per 1000	**263 fewer per 1000** (from 420 fewer to 0 more)
MBCT +mADM vs MBCT	249 (1 study)	⊕⊖⊖⊖ **very low**[3,4] due to risk of bias, imprecision	**RR 0.86** (0.74 to 0.99)	**Study population**	
				820 per 1000	**115 fewer per 1000** (from 8 fewer to 213 fewer)
				Moderate	
				820 per 1000	**115 fewer per 1000** (from 8 fewer to 213 fewer)

Notes:
[1] ROB high or unclear across 1-2 domains
[2] 95% CI crosses two clinical decision thresholds
[3] ROB high or unclear across multiple domains
[4] 95% CI crosses one clinical decision threshold

11.4 Economic evidence

11.4.1 Economic literature review

The systematic search of the literature identified 2 UK studies assessing the cost effectiveness of interventions aiming at preventing relapse in adults with depression (Kuyken et al., 2008 & 2015). Details on the methods used for the systematic search of the economic literature, including inclusion criteria for each review question, are described in Chapter 3. Full references and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix Q. Completed methodology checklists of the studies are provided in Appendix P. 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 R.

Both economic studies included in the review (Kuyken et al., 2008 & 2015) were conducted alongside RCTs (Kuyken2008, N=123; Kuyken2015, N=424) and assessed the cost effectiveness of mindfulness-based cognitive therapy (MBCT) with support to taper or discontinue antidepressant treatment versus maintenance antidepressant treatment plus medication adherence monitoring, in adults with at least 3 previous major depressive episodes, who were either in full or partial remission from their most recent depressive episode and on a therapeutic dose of maintenance antidepressants. The perspective of both analyses was the NHS and PSS; a broader societal perspective that included productivity losses and service user expenses was considered in a sensitivity analysis. Healthcare costs included intervention costs (provision of MBCT, medication, including support to taper or adhere to medication, hospital services (inpatient, outpatient, emergency department) and community health and social services (e.g., primary care by GPs, nurses and other healthcare professionals such as community psychiatrists and psychologists, social work, complementary therapies). National unit costs were used. Both studies used the percentage of people relapsing as measure of outcome; in addition, Kuyken and colleagues (2015) used QALYs based on EQ-5D (UK tariff) as a secondary outcome. The duration of the analyses ranged from 15 months (Kuyken et al., 2008) to 2 years (Kuyken et al., 2015).

Update 2017

Kuyken and colleagues (2008) reported that MBCT was more costly and more effective than maintenance antidepressant treatment, with an ICER of £335/additional relapse/recurrence prevented under a NHS and PSS perspective (figure converted from 2006 international dollars and uplifted to 2015 British pounds). As QALYs were not used as an outcome measure, the results of this study are not directly interpretable regarding the cost effectiveness of MBCT, as they require a judgement as to whether the extra benefit (prevention of one extra relapse) is worth the additional cost of £335. The study is thus only partially applicable to the NICE decision-making context and is characterised by minor limitations.

In the other study (Kuyken et al., 2015) MBCT was also more costly than maintenance antidepressant treatment and prevented a higher number of relapses, resulting in an ICER of £5,141 per relapse/recurrence averted under a NHS and PSS perspective (2015 prices). MBCT produced a lower number of QALYs compared with maintenance antidepressant treatment; therefore, based on the QALY outcome, MBCT does not appear to be cost-effective compared with maintenance antidepressant treatment as it is more costly and less effective. The study is directly applicable to the NICE decision-making context and is characterised by minor limitations.

11.4.2 Primary economic modelling

A decision-analytic model was developed to assess the relative cost effectiveness of pharmacological, psychological and combined interventions aimed at preventing relapse in people with depression that is in remission. The objective of economic modelling, the

methodology adopted, the results and the conclusions from this economic analysis are described in detail in Chapter 13. This section provides a summary of the methods employed and the results of the economic analysis.

Overview of economic modelling methods

A Markov model with a time horizon of 10 years was constructed to evaluate the relative cost effectiveness of a number of pharmacological, psychological and combined interventions for adults with depression that is in remission who are treated primarily in primary care. The economic analysis considered two different broad populations according to their risk of relapse as determined by the number of previous depressive episodes: adults with depression at medium risk of relapse (1-2 previous depressive episodes) and those at high risk of relapse (3+ previous depressive episodes). In those at medium risk of relapse, future depressive episodes were assumed to be less severe; in those at high risk of relapse, future depressive episodes were assumed to be more severe. These assumptions were based on GC expert advice, and aimed to cover a range of adults with depression that is in remission presenting in routine clinical practice The economic analysis considered separately populations that remitted following acute pharmacological, psychological and combination treatments. The time horizon (10 years) was selected to allow assessment of longer-term costs and benefits associated with relapse prevention treatment without introducing high complexity in the model structure. Based on the available evidence, the following analyses were carried out:

- Cost effectiveness of maintenance treatment with antidepressants versus clinical management with antidepressant tapering (reflected in pill placebo trial arms) in people at medium risk of relapse who remitted following acute pharmacological treatment and who experienced less severe depression if they relapsed; 4 analyses were undertaken that were specific to people who remitted following acute treatment with SSRIs, SNRIs, TCAs and mirtazapine.
- Cost effectiveness of maintenance treatment with antidepressants, MBCT plus clinical management with antidepressant tapering, MBCT combined with antidepressants, and clinical management with antidepressant tapering alone, in people at high risk of relapse who remitted following acute pharmacological treatment and who experienced more severe depression if they relapsed; group CT combined with antidepressants was added as an option in sensitivity analysis.
- Cost effectiveness of maintenance treatment with CT, fluoxetine, clinical management and no treatment in people at medium risk of relapse who remitted following acute psychological treatment and who experienced less severe depression if they relapsed.
- Cost effectiveness of maintenance treatment with CT, fluoxetine, clinical management and no treatment in people at high risk of relapse who remitted following acute psychological treatment and who experienced more severe depression if they relapsed; MBCT and group CT were added as options in sensitivity analysis.
- Cost effectiveness of maintenance treatment with combined pharmacological (fluoxetine) and psychological (CBT) intervention, pharmacological intervention alone (fluoxetine), psychological intervention plus clinical management with antidepressant tapering, and clinical management with antidepressant tapering alone in people at high risk of relapse who remitted following acute combination treatment and who experienced more severe depression if they relapsed.

The model structure considered the events of relapse (depressive episode), remission, and death. The probability of remission following a depressive episode was dependent on the time people spent in the depressive episode and was reduced as the time spent in the depressive episode increased. The probability of relapse for people in remission was dependent on the time people spent in remission and was reduced as the time spent in remission increased. Moreover, the risk of relapse depended on the number of previous episodes people had had in the past and increased with every new depressive episode that

Update 2017

was experienced. People receiving antidepressant treatment were at risk of developing common side effects from treatment. People in a depressive episode were assumed to be at increased mortality risk due to depression.

Efficacy data were derived from the guideline systematic review and were synthesised in a network meta-analysis (NMA). Baseline parameters (baseline risk of relapse) as well as the probability of recovery were estimated based on a review of naturalistic studies. The measure of outcome of the economic analysis was the number of QALYs gained. Utility data were derived from a systematic review of the literature, and were generated using EQ-5D measurements and the UK population tariff. The perspective of the analysis was that of health and personal social care services. Resource use was based on published literature, national statistics and, where evidence was lacking, the GC expert opinion. National UK unit costs were used. The cost year was 2016. Model input parameters were synthesised in a probabilistic analysis. This approach allowed more comprehensive consideration of the uncertainty characterising the input parameters and captured the non-linearity characterising the economic model structure. A number of one-way deterministic sensitivity analyses were also carried out.

Results are presented in the form of Incremental Cost Effectiveness Ratios (ICERs) following the principles of incremental analysis. Net Monetary Benefits (NMBs) are also provided. Results of probabilistic analysis have been summarised in the form of cost effectiveness acceptability curves (CEACs), which express the probability of each intervention being cost effective at different levels of willingness-to-pay per QALY gained (that is, at various cost effectiveness thresholds).

Overview of economic modelling results and conclusions

In people at medium risk of relapse who have remitted following acute pharmacological treatment (SSRIs, SNRIs, TCAs or mirtazapine) and who are expected to experience less severe depression if they relapse, maintenance pharmacological treatment is highly unlikely to be cost-effective compared with clinical management plus antidepressant drug tapering (probability of drugs being cost-effective ranging from 0.07 for SNRIs to 0.30 for SSRIs at the NICE lower cost-effectiveness threshold of £20,000/QALY). Maintenance pharmacological treatment, in particular with SSRIs, appears to be cost-effective if future episodes are more severe and as the risk of relapse increases (reflected in a higher number of previous episodes). This finding is explained by the low benefit-to-harm ratio of antidepressants in this population: the absolute risk of relapse is low (0.103 in the first year in people with one previous episode without maintenance drug treatment), the deterioration in HRQoL due to future relapse is milder (as relapses are less severe), and the risk of developing common side effects due to antidepressants and thus experiencing a utility decrement is relatively high (ranging from 0.117 with SSRIs to 0.163 with mirtazapine). However, as the number of previous episodes increases, the absolute risk of relapse increases and the preventive effect of maintenance drug treatment is enhanced; moreover, if relapses are more severe, the decrement in HRQoL resulting from each relapse increases, and the preventive effect of drugs has a larger (positive) impact on HRQoL. Consequently, the harms of maintenance drug treatment (side effects) are offset by its benefits (reduction in the number of relapses and larger improvement in HRQoL from prevention of relapses).

In people at high risk of relapse who have remitted following acute pharmacological treatment and who are expected to experience more severe depression if they relapse, the combination of MBCT with clinical management (antidepressant drug tapering) appears to be the most cost-effective option with quite high certainty (probability of being cost-effective 0.83 at the NICE lower cost-effectiveness threshold of £20,000/QALY). MBCT combined with antidepressant treatment is the second most cost-effective treatment option, followed by maintenance antidepressant treatment. The combination of MBCT with clinical management (antidepressant drug tapering) remained cost-effective in sensitivity analysis that included maintenance group CT combined with antidepressant treatment. MBCT plus clinical

Update 2017

management (antidepressant drug tapering) appeared to be the most cost-effective option under a range of scenarios explored in sensitivity analysis. However, if the preventive effect of MBCT lasts only one year, then the combination of MBCT plus antidepressant treatment becomes the most cost-effective intervention, followed by MBCT plus clinical management (antidepressant tapering), then antidepressant treatment alone, and, finally, clinical management and antidepressant drug tapering. Results are driven by the effectiveness of MBCT combined with the low intervention cost of (group-delivered) MBCT.

In people at medium risk of relapse who have remitted following acute psychological treatment and who are expected to experience less severe depression if they relapse, clinical management appears to be the most cost-effective intervention (with a probability of 0.58 at the NICE lower cost-effectiveness threshold of £20,000/QALY), followed by no treatment. Maintenance psychological treatment (CT) consisting of 10 individual hourly sessions appears to be the third most cost-effective option among those assessed in this analysis. However, if the preventive effect of CT can be achieved with 4 individual hourly sessions so that the intervention cost is greatly reduced, then CT appears to become the most cost-effective maintenance treatment option among those assessed in this population, provided that its relapse preventive effect lasts two years. The results are driven by the uncertainty characterising the clinical efficacy model input parameters, the relatively high cost of individual CT and the relatively low risk of relapse characterising the study population.

In people at high risk of relapse who have remitted following acute psychological treatment and who are expected to experience more severe depression if they relapse, clinical management appears to be the most cost-effective option (with a probability of 0.39 at the NICE lower cost-effectiveness threshold of £20,000/QALY) followed by maintenance CT. In sensitivity analysis that included group CT and MBCT, MBCT became the most cost-effective option, while group CT was the fourth most cost-effective option behind clinical management and maintenance CT. If the preventive effect of individual CT can be achieved with 4 hourly sessions, then CT becomes the most cost-effective option among all interventions assessed (including MBCT and group CT), even if its relapse preventive effect lasts only one year. The results are driven by the uncertainty characterising the clinical efficacy model input parameters and the relatively high cost of individual CT.

In people at high risk of relapse who have remitted following combined pharmacological and psychological acute treatment and who are expected to experience more severe depression if they relapse, maintenance pharmacological treatment alone appears to be the most cost-effective intervention followed by combination therapy. The probability of pharmacological treatment alone being the most cost-effective maintenance treatment option in this population is very high (0.95 at the NICE lower cost-effectiveness threshold of £20,000/QALY). It is noted that combination therapy is the most effective intervention; however, it has also a high intervention cost, mainly driven by the cost of maintenance psychological therapy, which comprises 10 individual CBT sessions. Nevertheless, even if the preventive effect of combined pharmacological and psychological therapy can be achieved with 4 individually delivered hourly sessions of CBT, meaning that the cost of combination therapy is greatly reduced, maintenance pharmacological treatment remains the most cost-effective treatment option. According to threshold analysis, combination therapy becomes the most cost-effective option when the psychological treatment component consists of 4 individual hourly sessions, and the population has at least 7 previous depressive episodes, so that the risk of relapse is increased and the impact of the preventive effect of combination therapy is enhanced. Psychological therapy plus clinical management (antidepressant drug tapering) appears to be marginally less cost-effective than clinical management (drug tapering) alone; its relative cost effectiveness versus clinical management improves when the number of previous episodes (and therefore the risk of future relapses) increases and when psychological therapy comprises 4 individual sessions (rather than 10). Results are driven by the high effectiveness of antidepressant therapy along or in combination with psychological therapy and the high cost of psychological therapy if it consists of 10 individual CBT sessions.

Update 2017

Results of the economic analysis were overall robust to different scenarios explored through sensitivity analysis. In general, the relative cost effectiveness of more effective interventions improved when the risk of relapse (as reflected in number of previous episodes) increased, because their preventive effect had a greater impact (as a higher number of future relapses was avoided), and associated cost-savings offset the maintenance intervention costs. The cost effectiveness of individual psychological interventions improved when the number of sessions was reduced, provided that their relapse preventive effect was fully retained.

Conclusions from the guideline economic analysis refer mainly to people with depression who are predominantly treated in primary care; however, they may be relevant to people in secondary care as well, especially given that clinical evidence was derived almost exclusively from studies conducted in secondary care settings (however, it needs to be noted that costs utilised in the guideline economic model were mostly relevant to primary care).

11.5 Clinical evidence statements

11.5.1 Psychological interventions versus control

- Low quality evidence from 3-4 RCTs (k=3-4, n=426-471) showed that at 12 month follow-up individuals treated with CBT or CT experienced a significant reduction in risk of relapse compared with those who had received a control treatment, however by 24 month follow-up the advantage in the CBT/CT group remained statistically significant but was no longer clinically important.
- Low-moderate quality evidence from 2-9 RCTs (k=2-9, n=627-1000) showed that at 12 month follow-up individuals treated with MBCT had a statistically significant but not clinically important reduction in risk of relapse compared with those who had received a control treatment, however this advantage had disappeared by 24 month follow-up.
- Very low quality evidence from 3 RCTs (k=3, n=187-193) showed no reduced risk of relapse in individuals treated with IPT over control at either 12 or 24 month follow-up.
- Very low quality evidence from 1 RCT (k=1, n=82) showed significantly reduced risk of relapse in patients treated with CBASP over control at 12 month follow-up.

11.5.2 Psychological interventions versus psychological interventions

- Low quality evidence from 1 RCT (k=1, n=180) showed no difference in risk of relapse at 12 month follow-up between patients treated with CBT or psychoeducation.
- Very low-low quality evidence from 1 RCT (k=1, n=87-88) showed no difference in risk of relapse between different frequencies of IPT sessions (weekly, bi-monthly or monthly) at 24 month follow-up.

11.5.3 Psychological versus pharmacological interventions

- Low-very low quality evidence from 2 RCTs (k=2, n=155-172) showed no difference in relapse rate between patients treated with CBT or an antidepressant at either 12 or 24 month follow-up.
- Low quality evidence from 2 RCTs (k=2, n=115) showed a clinically important but not statistically significant increase in risk of relapse in patients treated with an antidepressant when compared with those treated with IPT.

11.5.4 Pharmacological monotherapy compared with control

- Very low quality evidence from 48 RCTs (k=48, n=9105) showed a reduced risk of relapse in patients treated with an antidepressant versus placebo. This effect held across all included drug classes; SSRIs (k=22, n=4360; including sertraline, fluoxetine, fluvoxamine, escitalopram, citalopram, paroxetine and paroxetine plus lithium or desipramine), TCAs (k=11, n=1299), SNRIs (k=6, n=1450), and 'other' antidepressants (k=5, n=829).

Update 2017

- Low quality evidence from 1 RCT (k=1, n=771) showed a reduced risk of relapse in patients treated with quetiapine compared with placebo.
- Very low quality evidence from 1 RCT (k=1, n=107) showed a reduced risk of relapse in patients treated with amitriptyline compared with placebo.

11.5.5 Full dose versus reduced dose pharmacological treatment

- Very low-low quality evidence from 2 RCTs (N=835) showed a clinically important but not statistically significant reduction in risk of relapse in patients treated with a full dose compared with a half-dose of an SSRI, but not in those treated with a TCA or an antidepressant overall.

11.5.6 Pharmacological augmentation strategies

- Low quality evidence from 3 RCTs (k=3, n=160) showed a clinically important but not statistically significant reduction in risk of relapse in patients whose antidepressant therapy was augmented with lithium, compared with placebo.
- Low quality evidence from 1 RCT (k=1, n=241) showed no benefit from augmenting antidepressant treatment with risperidone compared with placebo.

11.5.7 Pharmacological interventions in ECT responders

- Very low quality evidence from 1 RCT (k=1, n=39) showed no difference in risk of relapse between patients treated with fluoxetine plus placebo or fluoxetine plus melatonin at 12 week follow-up.
- Very low-low quality evidence from 2 different RCTs (k=2, n=57-201) showed no difference in risk of relapse at 6 months in patients treated with ECT compared with nortriptyline plus lithium or nortriptyline plus lithium compared with nortriptyline alone, but a reduced risk of relapse in those treated with nortriptyline plus lithium compared with placebo.
- Very low-low quality evidence from 2 RCTs (k=2, n=27-56) showed a reduced risk of relapse in patients treated with imipramine but not nortriptyline when compared with placebo at 6 month follow-up.
- Very low quality evidence from 1 RCT (k=1, n=43) showed a clinically important but not statistically significant reduction in risk of relapse in patients treated with paroxetine compared with imipramine at 6 month follow-up.
- Low quality evidence from 1 RCT (k=1, n=56) showed a reduced risk of relapse in patients treated with a combination of ECT and an antidepressant compared with an antidepressant alone at 12 month follow-up.

11.5.8 Combined psychological and pharmacological interventions versus pharmacological interventions

- Very low quality evidence from 1 RCT (k=1, n=53) showed a clinically important but not statistically significant reduced risk of relapse in patients treated with a combination of imipramine and IPT versus imipramine alone.
- Very low quality evidence from 1 RCT (k=1, n=68) showed no difference in risk of relapse between patients treated with MBCT plus an antidepressant or an antidepressant alone.
- Very low quality evidence from 1 RCT (k=1, n=63) showed no difference in risk of relapse between patients treated with a combination of paroxetine and IPT versus paroxetine alone.
- Very low quality evidence from 2 RCTs (k=2, n=177) showed no difference overall in the risk of relapse between patients treated with a combination of CBT or CT and an antidepressant versus an antidepressant alone, or specifically for the combination of CT plus fluoxetine versus fluoxetine alone (k=1, n=132), however there was a clinically

Update 2017

important but not statistically significant reduction in risk of relapse in patients treated with CBT plus any antidepressant versus any antidepressant versus any antidepressant alone (k=1, n=45).

11.64 Economic evidence statements

- Evidence from a single UK study conducted alongside a RCT (N =424) suggests that MBCT is not cost-effective compared with maintenance antidepressant treatment in people who have had at least 3 previous depressive episodes and are in full or partial remission from their most recent episode following acute pharmacological treatment. The study is directly applicable to the NICE decision-making context and is characterised by minor limitations. Evidence from another UK study conducted alongside a RCT on the same population (N=123) is inconclusive regarding the cost effectiveness of MBCT compared with maintenance antidepressant treatment, as the outcome measure was not the QALY and interpretation of the results depends on the willingness to pay in order to avoid an additional relapse/recurrence of depression. Therefore the study, although it was conducted in the UK, is only partially applicable to the NICE decision-making context. The study is characterised by minor limitations.
- Evidence from the guideline economic modelling suggests that in people at medium risk of relapse who have remitted following acute pharmacological treatment and who are expected to experience less severe depression if they relapse, maintenance pharmacological treatment with the same drug they had received as acute treatment over 2 years is not cost-effective versus clinical management (antidepressant tapering) due to the high harm-to-benefit ratio of maintenance drug treatment in this population. The cost effectiveness of maintenance drug treatment increases as the severity of depression increases and as the risk for future relapses, as determined by the number of previous episodes, increases. This evidence refers mainly to people treated in primary care; however, it may be relevant to people treated in secondary care as well, given that the vast majority of clinical evidence was derived from secondary care settings. The analysis is directly applicable to the NICE decision-making context and is characterised by minor limitations.
- Evidence from the guideline economic modelling suggests that in people at high risk of relapse who have remitted following acute pharmacological treatment and who are expected to experience more severe depression if they relapse, maintenance treatment with MBCT in combination with clinical management (antidepressant drug tapering) is the most cost-effective option with high certainty, followed by combination of MBCT with antidepressant treatment. Maintenance antidepressant treatment alone is more cost-effective than clinical management with antidepressant tapering. If the preventive effect of MBCT lasts only one year, then the combination of MBCT plus antidepressant treatment becomes the most cost-effective intervention, followed by MBCT plus clinical management (antidepressant tapering). This evidence refers mainly to people treated in primary care; however, it may be relevant to people treated in secondary care as well, given that the vast majority of clinical evidence was derived from secondary care settings. The analysis is directly applicable to the NICE decision-making context and is characterised by minor limitations.
- Evidence from the guideline economic modelling suggests that in people at medium risk of relapse who have remitted following acute psychological treatment and who are expected to experience less severe depression if they relapse, maintenance high intensity CT (comprising 10 individual hourly sessions) is unlikely to be cost-effective, and clinical management or no treatment should be preferred instead. However, if the preventive effect of CT can be achieved with 4 individual hourly sessions so that the intervention cost is greatly reduced, then maintenance CT is potentially cost-effective provided that its relapse preventive effect lasts two years. This evidence refers mainly to people treated in primary care; however, it may be relevant to people treated in secondary care as well, given that the vast majority of clinical evidence was derived from secondary care settings.

Update 2017

The analysis is directly applicable to the NICE decision-making context and is characterised by minor limitations.

- Evidence from the guideline economic modelling suggests that in people at high risk of relapse who have remitted following acute psychological treatment and who are expected to experience more severe depression if they relapse, maintenance CT comprising 10 individual hourly sessions and with an effect that lasts two years is marginally less cost-effective than clinical management. Maintenance CT consisting of 4 individual hourly sessions (provided that it can achieve the same effect as CT comprising 10 individual sessions over a minimum of one year) is more cost-effective than clinical management. MBCT also appears to be a cost-effective option for this population, although less cost-effective than 4 individual hourly sessions of CT (provided that its effect is equal to that of CT comprising 10 individual sessions). This evidence refers mainly to people treated in primary care; however, it may be relevant to people treated in secondary care as well, given that the vast majority of clinical evidence was derived from secondary care settings. The analysis is directly applicable to the NICE decision-making context and is characterised by minor limitations.
- Evidence from the guideline economic modelling suggests that in people at high risk of relapse who have remitted following combined pharmacological and individual psychological acute treatment and who are expected to experience more severe depression, maintenance pharmacological treatment alone is highly likely the most cost-effective treatment option. Combination therapy is the most cost-effective option if it includes a less intensive psychological component (e.g. 4 individual hourly sessions that retain the effect of 10 sessions), and the population's risk of relapse is quite high, as determined by a higher number (at least 7) of previous depressive episodes. Maintenance individual psychological therapy plus clinical management (drug tapering) becomes potentially more cost-effective than clinical management alone as the number of previous episodes (and thus the risk of relapse characterising the study population) increases or if the number of individual sessions is reduced to 4 (provided that the effect of 10 individual sessions can be achieved for a minimum of one year). This evidence refers mainly to people treated in primary care; however, it may be relevant to people treated in secondary care as well, given that the vast majority of clinical evidence was derived from secondary care settings. The analysis is directly applicable to the NICE decision-making context and is characterised by minor limitations.

Update 2017

11.7 From evidence to recommendations

11.7.1 Relative values of different outcomes

The outcome of interest to the GC for this review was relapse. The time points of interest for psychological interventions were 12 and 24 months, and for pharmacological interventions were endpoint and follow-up, which varied by study.

11.7.2 Trade-off between clinical benefits and harms

Psychological interventions

The evidence for psychological interventions to reduce risk of relapse was predominantly of very low quality, although some evidence was of moderate quality, and was generally from few trials of fairly small numbers of patients. The most significant evidence came from a small trial of a fairly rarely utilised psychological intervention (CBASP). However the GC were concerned about the generalisability of this trial and were not sufficiently confident in the findings to make a recommendation for its use. The evidence seemed to show no particular benefit from providing combination treatment over pharmacological treatment, and suggested that by 24 month follow-up there was no difference between people treated with CBT and those treated with antidepressants, and that antidepressants may be worse than IPT for

reducing relapse. In those patients who had only partially remitted before entering relapse prevention, CBT and CT were highly effective over an extended time-period. Additionally in terms of absolute numbers, rather than purely 'clinical importance' as defined a priori, the GC agreed that CBT and MBCT appeared to be beneficial over a substantial follow-up period, were the most effective interventions available, and had no untoward side effects. Furthermore CBT is recommended as first line treatment for depression and so the GC decided to recommend that treatments such as CBT, which have an explicit relapse prevention component as part of the core treatment, should be offered to people with less severe depression who are at risk of relapse.

The psychological interventions included in this review included a range of different treatment approaches. Some models consisted of a relapse focused extension of standard treatment, typically 3-4 sessions over a 2 month period. The GC used this information to further develop the recommendation for relapse prevention in people with less severe depression.

Pharmacological interventions

The evidence for pharmacological interventions was generally of very low to low quality, but came from many trials with large numbers of participants. Antidepressants appeared to be an effective relapse prevention treatment, although dose reduction did not appear to impact upon the effectiveness of antidepressants. The evidence for lithium augmentation was of low quality and very limited, and showed only a trend for benefit. In people who had responded to ECT, a combination of ECT and an antidepressant appeared more effective by 12 month follow-up than receiving an antidepressant alone. The evidence for those who had responded to ECT, however, was very limited and mixed, and the GC had insufficient confidence in the evidence to make any firm recommendations.

The GC were aware of the limitations in the evidence in support of medication but recognised that it did appear to be effective and also that some people would not wish to take up the offer of a psychological intervention. They therefore recommended that medication be offered for relapse prevention in combination with psychological interventions but also on its own if a person did not want psychological, intervention.

The GC considered the clinical benefit in this instance to be a reduced risk of relapse. The harms were considered to be an increased risk of relapse, the provision of ineffective treatments, or people having side effects that may impact negatively upon quality of life or decrease engagement with treatment, potentially in itself inducing a relapse.

The GC discussed the issue of patients remaining on pharmacotherapy when no further benefit may be obtained, potentially with debilitating adverse effects, and for this reason they recommended that follow-up be regular, and the period between reviews no more than 12 months.

ECT

The review of ECT for the updated guideline found relatively little additional data to update the reviews undertaken for the original NICE TA (NICE, 2003) and the revision of the guideline in 2009 (NICE, 2009). The focus of the new evidence was on mode of administration e.g. unilateral versus bilateral. As in 2009 the GC came to the conclusion that bilateral ECT is more effective than unilateral for people with depression. Having carefully considered the evidence the GC did not think that the evidence reviewed for this guideline supported or required any changes to the existing recommendations. ECT primarily remains a treatment for severe often life threatening depression.

For cognitive impairment, it remains unclear to what degree the trade-off between efficacy and cognitive side effects can be avoided by manipulating dose and electrode placement. There is, however, evidence that bilateral ECT causes more cognitive impairment than unilateral ECT and that the cognitive impairment and efficacy from unilateral ECT are dose

Update 2017

related. The data on continuation/maintenance ECT that support at least equal efficacy in preventing relapse compared with pharmacotherapy remains limited. Systematic, prospective assessment of longer-term cognitive effects of continuation/maintenance ECT are also limited although those available do not suggest cumulative cognitive adverse effects. Given the relative lack of data, the GC again made no treatment recommendation regarding continuation/maintenance ECT.

11.7.37 Trade-off between net health benefits and resource use

The guideline economic analysis showed that in people at medium risk of relapse who have remitted following acute pharmacological treatment and who are expected to experience less severe depression if they relapse, maintenance pharmacological treatment is not cost-effective versus clinical management (antidepressant tapering) due to the high harm-to-benefit ratio of maintenance drug treatment in this population. However, the analysis showed that the cost effectiveness of maintenance drug treatment improves as the severity of depression increases and as the risk for future relapses increases.

In people at high risk of relapse who have remitted following acute pharmacological treatment and who are expected to experience more severe depression if they relapse, maintenance treatment with MBCT in combination with clinical management (antidepressant drug tapering) appeared to be the most cost-effective option with high certainty, followed by the combination of MBCT with antidepressant treatment. Maintenance antidepressant treatment alone is more cost-effective than clinical management with antidepressant tapering. If the preventive effect of MBCT lasts only one year, then the combination of MBCT plus antidepressant treatment appears to be the most cost-effective intervention, followed by MBCT plus clinical management (antidepressant tapering).

The GC noted that evidence from a RCT conducted in the UK was inconsistent with the results of the guideline economic modelling, as it suggested that MBCT was not cost-effective compared with maintenance antidepressant treatment in people at high risk of relapse (at least 3 previous depressive episodes) who were in full or partial remission from their most recent depressive episode following acute drug treatment. In this study, MBCT reduced the risk of relapse relative to maintenance antidepressant treatment, so it was more effective in this aspect, but also resulted in a lower number of QALYs, which was a rather unexpected finding, as a reduced risk of relapse is expected to be associated with longer periods of remission and, subsequently, a higher HRQoL. In contrast, the guideline economic model, which attached a higher utility value in the health state of remission than in the health state of relapse, found a better effect of MBCT compared with maintenance antidepressant treatment regarding relapse prevention, and, consequently, a higher gain in QALYs. In addition, the economic model had a longer time horizon compared with this RCT's duration of follow-up, which may also contribute to the discrepancy of findings between the guideline economic analysis and the analysis conducted alongside the RCT.

In another RCT conducted in the UK on the same population, evidence was inconclusive regarding the cost effectiveness of MBCT compared with maintenance antidepressant treatment, as the outcome measure was not the QALY and interpretation of the results required judgements on the value of preventing an additional relapse/recurrence of depression. Nevertheless, in this analysis MBCT was more effective in preventing relapses than maintenance antidepressant treatment, which is consistent with the findings of the guideline economic analysis.

In people at medium risk of relapse who have remitted following acute psychological treatment and who are expected to experience less severe depression if they relapse, the guideline economic analysis suggested that maintenance high intensity CT (comprising 10 individual hourly sessions) was unlikely to be cost-effective, and clinical management or no treatment should be preferred instead. However, if the preventive effect of CT can be achieved with 4 individual hourly sessions so that the intervention cost is greatly reduced,

Update 2017

then maintenance CT is potentially cost-effective provided that its relapse preventive effect lasts two years. The GC considered 10 sessions of psychological therapy to be unrealistically high as maintenance treatment, and expressed the view that 4 sessions are adequate to maintain a relapse preventive effect.

In people at high risk of relapse who have remitted following acute psychological treatment and who are expected to experience more severe depression if they relapse, maintenance CT comprising 10 individual hourly sessions and with an effect that lasts two years was marginally less cost-effective than clinical management. On the other hand, maintenance CT consisting of 4 individual hourly sessions (provided that it can achieve the same effect as CT comprising 10 individual sessions over a minimum of one year) was shown to be more cost-effective than clinical management. MBCT also appeared to be a cost-effective option for this population in the guideline economic analysis, although less cost-effective than 4 individual hourly sessions of CT.

In people at high risk of relapse who have remitted following combined pharmacological and individual psychological acute treatment and who are expected to experience more severe depression, the economic analysis showed that maintenance pharmacological treatment alone was highly likely the most cost-effective treatment option. Combination therapy is the most cost-effective option if it includes a less intensive psychological component (e.g. 4 individual hourly sessions that retain the effect of 10 sessions), and the population's risk of relapse is quite high, as determined by a higher number (at least 7) of previous depressive episodes. Maintenance individual psychological therapy plus clinical management (drug tapering) becomes potentially more cost-effective than clinical management alone as the number of previous episodes increases or if the number of individual sessions is reduced to 4 (provided that the effect of 10 individual sessions can be achieved for a minimum of one year).

The guideline economic modelling considered predominantly people treated in primary care; however, the GC noted that the vast majority of clinical evidence was derived from secondary care settings, due to lack of relevant evidence derived from primary care settings. The GC considered it reasonable and essential to extrapolate the secondary care evidence to the primary care population when formulating recommendations due to a lack of more relevant evidence.

The GC noted that the definition of 'medium' and 'high' risk of relapse in the economic analysis was based exclusively on the number of previous depressive episodes experienced by the study population (1-2 previous episodes and 3+ previous episodes, respectively) and was made for practical reasons, in order to populate the economic model. However, it was acknowledged that the risk of future relapse is determined by a combination of several other factors, including the frequency of previous depressive episodes and how recently these were experienced; the presence of residual symptoms and unhelpful coping styles such as avoidance and rumination; the severity of previous episodes and the presence of functional impairment and risk-to-self during the episodes; the effectiveness of previous interventions for treatment and relapse prevention; the presence of other chronic physical health or mental health problems and the presence of personal, social and environmental factors. Therefore, the population at a 'higher' risk of relapse in clinical practice may include people with 1-2 previous episodes (considered as being at 'medium' risk in the economic analysis) if other factors increasing the risk of relapse are present.

The GC reviewed the results of the guideline economic analysis and noted that in people at medium risk of relapse, defined as having had 1-2 previous depressive episodes, relapse preventive interventions were not cost-effective compared with clinical management (and drug tapering, if relevant). However, as expected, the cost effectiveness of relapse preventive interventions improves as the severity of depression increases and as the risk for future relapses grows, as in both cases there is more scope for gains in HRQoL if relapses are prevented. A range of relapse preventive interventions were cost-effective in people with

Update 2017

depression that was in remission and who were at high risk of relapse, defined as having had at least 3 previous depressive episodes, depending on the acute treatment that led to remission of the episode.

Therefore the GC decided to recommend cost-effective interventions, as identified in the guideline economic analysis, for people at a 'higher' risk of relapse, which should be estimated after considering all the factors affecting the risk of relapse, and not based solely on the number of previous depressive episodes. The GC did not make recommendations specifically for people at 'low' risk of relapse, as relapse preventive interventions are not cost-effective in this population and, for maintenance antidepressant treatment, harms (side effects) could potentially outweigh benefits (limited scope for prevention of new depressive episodes in a population with a low baseline risk of relapse).

The GC considered that the relative low cost of administration of ECT and its potential benefit in severe depression did not represent a significant cost impact given the very low numbers of people who receive ECT and the potential savings that might accrue in terms of reduced length of hospital admissions.

11.7.4 Quality of evidence

The quality of the evidence was assessed using GRADE. The GC noted generally that the evidence for psychological interventions was much longer-term than for pharmacological interventions. The largest body of evidence for pharmacological evidence had a follow-up of only 6 months, compared with follow-up of at least 12 months and in many cases 24 months for psychological interventions.

The GC noted the lack of data from the primary care population and agreed to recommend further research to establish what the rate of relapse is in people with depression who present, and are treated, in primary and secondary care.

The GC also recognised that there was limited data comparing psychological interventions for relapse against each other and against antidepressants. They therefore recommended further research in this area.

11.7.5 Other considerations

The GC discussed the importance of understanding that a relapse was a possibility. The lay members on the GC explained that it can be quite empowering to understand that depression can be a recurrent condition, and that a relapse does not indicate any kind of failure on the part of the person with depression. Therefore the GC agreed that it would be helpful to recommend that this possibility is discussed at an appropriate time.

The GC also considered the issue of patient choice, and the need to factor this into any recommendations. For this reason they incorporated different combinations of options into the recommendations addressing the possibility that someone who has remitted following pharmacotherapy may not wish to continue with this, or vice versa. Additionally they discussed the importance of life events and stressors to the potential for relapse. This is not a factor than can be adequately captured by a systematic review, however the GC on the basis of their expert knowledge decided it was crucial that this be explicitly addressed within the recommendations.

11.8 Recommendations

96. Discuss the likelihood of having a relapse with people who have recovered from depression. Explain:

Update 2017

- that a history of previous relapse increases the chance of further relapses
- the potential benefits of relapse prevention. [new 2017]

97. Take into account that the following may increase the risk of relapse:

- how often a person has had episodes of depression, and how recently
- any other chronic physical and mental health problems
- any residual symptoms and unhelpful coping styles, for example avoidance and rumination)
- how severe their symptoms were, risk to self and if they had functional impairment in previous episodes of depression
- the effectiveness of previous interventions for treatment and relapse prevention
- personal, social and environmental factors. [new 2017]

98. For people who have recovered from less severe depression when treated with medication (alone or in combination with a psychological therapy), but are assessed as having a higher risk of relapse, consider:

- psychological therapy (CBT) with an explicit focus on relapse prevention, typically 3–4 sessions over 1–2 months
- continuing their medication. [new 2017]

99. For people who have recovered from more severe depression when treated with medication (alone or in combination with a psychological therapy), but who may be at higher risk of relapse, offer:

- a psychological therapy [mindfulness-based cognitive therapy (MBCT) or group CBT] in combination with medication, **or**
- psychological therapy (MBCT or group CBT) with a focus on relapse prevention if the person wants to stop taking medication. [new 2017]

100. For people who have recovered from depression when treated with a psychological therapy, but are assessed as having a higher risk of relapse, offer further psychological therapy (see recommendation 98). [new 2017]

101. For people who are continuing with medication to prevent relapse, maintain the same dose unless there is good reason to reduce it (such as adverse effects). [new 2017]

102. For people continuing with medication to prevent relapse, hold reviews at 3, 6 and 12 months after maintenance treatment has started. At each review:

- monitor mood state using a formal validated rating scale, for example the PHQ-9
- review side effects
- review any personal, social and environmental factors that may impact on the risk of relapse
- agree the timescale for further review (no more than 12 months). [new 2017]

103. At all further reviews for people continuing with antidepressant medication to prevent relapse:

Update 2017

- assess the risk of relapse
- discuss the need to continue with medication. [new 2017]

104. Offer group CBT (or MBCT for those who have had 3 or more previous episodes of depression) for preventing relapse to people who are assessed as being at higher risk of relapse and who recovered with medication but who want to stop taking it. [new 2017]

105. When choosing a psychological therapy for preventing relapse for people who recovered with initial psychological therapy, offer:

- 4 more sessions of the same treatment if it has an explicit relapse prevention component, or
- group CBT (or MBCT for those who have had 3 or more previous episodes of depression) if initial psychological therapy had no explicit relapse prevention component. [new 2017]

106. Re-assess a person's risk of relapse when they finish a psychological relapse prevention intervention. Discuss the need for continuing treatment with the person if necessary. [new 2017]

107. Deliver MBCT for people assessed as having a higher risk of relapse in groups of up to 15 participants. Meetings should last 2 hours once a week for 8 weeks, with 4 follow-up sessions in the 12 months after treatment ends. [new 2017]

108. Deliver group CBT for people assessed as having a higher risk of relapse in groups of up to 12 participants. Sessions should last 2 hours once a week for 8 weeks. [new 2017]

Update 2017

Electroconvulsive therapy

109. Consider electroconvulsive therapy (ECT) for acute treatment of more severe depression if:

- the more severe depression is life-threatening and a rapid response is needed, or
- multiple pharmacological and psychological treatments have failed. [2017]

110. For people whose depression has not responded well to ECT previously, only consider a repeat trial of ECT after:

- reviewing the adequacy of the previous treatment course
- considering all other options
- discussing the risks and benefits with the person or, if appropriate, their advocate or carer. [2017]

111. Make sure people with depression who are going to have ECT are fully informed of the risks, and with the risks and benefits specific to them. Take into account:

- the risks associated with a general anaesthetic
- any medical comorbidities
- potential adverse events, in particular cognitive impairment
- if the person is older, the possible increased risk associated with ECT treatment for this age group

- the risks associated with not having ECT.

Document the assessment. [2017]

112. Make the decision to use ECT together with the person with depression if they have the capacity to give consent. Take into account the requirements of the Mental Health Act 2007 (if applicable), and make sure:

- valid, informed consent is given without pressure or coercion from the circumstances or clinical setting
- the person is aware of their right to change their mind and withdraw consent at any time
- there is strict adherence to recognised guidelines on consent, and advocates or carers are involved to help informed discussions. [2017]

113. If a person with depression cannot give informed consent, only give ECT if it does not conflict with an advance treatment decision the person made [2017]

114. For a person with depression who is going to have ECT, assess their cognitive function:

- before the first treatment
- at least every 3–4 treatments
- at the end of the treatment course. [2017]

115. Check for the following in cognitive function assessments for people having ECT:

- orientation, and time to reorientation after each treatment
- measures of new learning, retrograde amnesia and subjective memory impairment, carried out at least 24 hours after a treatment. [2017]

116. If a person shows signs of significant cognitive impairment at any stage of ECT treatment, consider:

- changing from bilateral to unilateral electrode placement, or
- reducing the stimulus dose, or
- stopping treatment. [2017]

117. When making any decision to change ECT treatment, make sure:

- the person (or their family members, carers or advocate) still gives their valid, informed consent
- the person is fully aware of the risks and benefits of the treatment change. [2017]

118. When giving ECT to a person with depression:

- base the electrode placement and stimulus dose, related to seizure threshold, on a balance of effectiveness against the risk of cognitive impairment
- be aware that bilateral ECT is more effective than unilateral ECT, but may cause more cognitive impairment
- be aware that with unilateral ECT a higher stimulus dose can be more effective, but can also increase cognitive impairment. [2017]

Update 2017

119. Assess a person's clinical status after each ECT treatment using a formal valid outcome measure (HRDS or MDRAS). [2017]

120. Stop ECT treatment for a person with depression:

- straightaway, if the side effects outweigh the potential benefits, or
- when remission has been achieved. [2017]

121. If a person's depression has responded to a course of ECT:

- start (or continue) antidepressant medication to prevent relapse
- consider lithium augmentation of antidepressants. [2017]

11.9 Research recommendations

5. What is the rate of relapse in people with depression who present, and are treated, in primary and secondary care, and what factors are associated with increased risk of relapse?

Statement: A large-scale, long-term cohort study should be undertaken to establish the rate of relapse in adults with depression who are successfully treated in primary care and secondary care, and the factors associated with an increased risk of relapse in this population.

Rationale: The current understanding of the rate of relapse in depression is that it is high and may be up to 50% after a first episode, rising to 80% in people who have had three or more episodes of depression. However, most studies have been undertaken in the secondary care setting and whether these figures represent the actual rate of relapse in primary care populations is uncertain. In addition, beyond the number of previous episodes and the presence of residual symptoms, there is also considerable uncertainty about what other factors (biological, psychological or social) might be associated with an increased risk of relapse. This cohort study will enable clinicians to more accurately identify those at risk of relapse, and provide relapse prevention strategies for these individuals. Accordingly, this would improve clinical outcomes and quality of life in patients as well as facilitating more targeted use of NHS resources.

6. What is the comparative effectiveness and cost effectiveness of group based psychological treatments in preventing relapse in people with depression (compared to each other and antidepressant medication) for people who have had a successful course of treatment with antidepressants or psychological therapies?

Statement: A randomised controlled trial should be conducted to establish whether, in adults in remission from depression following either antidepressant treatment or psychological therapies, group CBT, MBCT or medication results in lower incidence of depressive relapse.

Rationale: Depressive relapse is a frequent occurrence with implications for the wellbeing and quality of life for the individual and financial implications for the NHS. Antidepressants can be effective in preventing relapse but not all service users can tolerate them or wish to take them long-term. Two, group based psychological interventions (group CBT and mindfulness based cognitive therapy) have been developed and shown to be effective primarily in trials when compared to treatment as usual. However, they have not been compared with each other and only in a limited way against antidepressants. The randomised controlled trial should be designed to identify both moderators and mediators of treatment effect, have a minimum follow up period of two years, assess any adverse events and the relative cost-effectiveness of the interventions and test for both superiority and equivalence.

Update 2017

12 Access to services

12.1 Introduction

Improving access to health and social care should help people get the resources they need in order to preserve and improve their health and well-being. Access is complex and depends on a range of factors such as adequacy of supply, uptake, effectiveness of services, and equity in meeting the needs of different groups (Gulliford, Figueroa-Munoz et al. 2002, Dixon-Woods, Kirk et al. 2005, NICE 2011). This chapter focuses particularly on uptake and equity issues. In terms of uptake, Lepine et al. (Lepine, Gastpar et al. 1997) found that only 57% of those diagnosed with depression sought help.

Equity of access to treatment is also a major concern. In the latest National Psychiatric Morbidity Survey, after controlling for need, white British women and people in mid-life (aged 35-54) were more likely to receive treatment for depression than people in Black/Black British ethnic groups (McManus, Bebbington et al. 2016). Public engagement findings by the Mental Health Taskforce (Mental-Health-Taskforce 2015) highlighted accessibility of services as an issue with people wanting improvements to target those experiencing the poorest access, experience and outcomes. Of course, when considering access, not all treatments are the same. Patients may prefer talking therapies to anti-depressant medication (Gaudiano and Miller 2013, McHugh, Whitton et al. 2013), but the evidence suggests that medication is offered much more commonly (McManus, Bebbington et al. 2016). A recent report suggested that 40% of people had to ask for psychological therapies rather than being offered them proactively (MIND 2013). Computerised CBT may appear to be an effective and convenient option for some people, but uptake appears low and dropout relatively high, with just over half of people completing a full course of treatment (Waller and Gilbody 2009).

Poor access to services may be a greater problem in some groups than others. The Guideline Committee chose to focus on uptake and equity of access for three groups whose needs may not have been adequately met based on previous evidence reviews (Dixon-Woods, Kirk et al. 2005, NICE 2011):

- older people (Crawford, Prince et al. 1998, Department-of-Health 2011, NAPT 2013),
- BME groups (Bhui, Bhugra et al. 2001, Bhui, Stansfeld et al. 2003, Suresh and Bhui 2006, Cooper, Spiers et al. 2013) and
- men (Shiels, Gabbay et al. 2004, Addis 2008, Martin, Neighbors et al. 2013, Stansfeld, Clark et al. 2016).

High levels of depression and low levels of service use have been reported among older adults from UK minority ethnic groups (Lawrence, Banerjee et al. 2006). GPs reported that older patients tended not to mention psychological difficulties, tending to see these as part of ageing (Murray, Banerjee et al. 2006). Older men were particularly reluctant and were more vulnerable to severe depression and suicide (Murray, Banerjee et al. 2006). Perceived stigma about having a mental health problem was seen as a barrier to seeking help.

People from BME backgrounds access help less often from their GPs for mental health problems than the white population. This has been found with people from Black Caribbean (Nazroo, Edwards et al. 1997) and South Asian (Anand and Cochrane 2005). They are also less likely to be diagnosed if they do consult (Odell, Surtees et al. 1997, Maginn, Boardman et al. 2004). Memon, Taylor et al. (2016) found that the relationship between service user and healthcare provider in people from BME backgrounds was affected by factors such as a lack of awareness of different services by service users and providers, language barriers, poor communication, an imbalance of power and authority, as well as cultural insensitivity. The study concluded that this patient group need support with mental health literacy and increased awareness of mental health conditions. Illness perceptions of depression may also affect help-seeking. Compared to white British women, Black African (Brown, Casey et al.

Update 2017

2011) and Indian women (Taylor, Brown et al. 2013) thought depression was less amenable to treatment.

Men consult less frequently than women for emotional problems, particularly for depression (Moller-Leimkuhler 2002, Prins, Verhaak et al. 2008). Different health beliefs appear relevant: men perceive less of a need for treatment (Edwards, Tinning et al. 2007) and have less confidence in mental health professionals (Kessler, Brown et al. 1981). Masculinity is also important in reducing help-seeking (Seidler, Dawes et al. 2016). House et al (submitted) found considerable shame is experienced by men who experience depressive problems. Suicide is strongly associated with mental illness, particularly depression and the male suicide rate in the UK is current three times the female rate (ONS 2016).

12.21 Review question

- In adults (18 years and older) at risk of depression or (anxiety disorders) from particular vulnerable groups (older people, BME groups and men) do service developments and interventions which are specifically designed to promote access, increase the proportion of people from the target group who access treatment, when compared with standard care?

The review protocol summary and the eligibility criteria used for this section of the guideline, can be found in Table 241. A complete list of review questions and review protocols can be found in Appendix F; further information about the search strategy can be found in Appendix H.

Table 241: Clinical review protocol summary for the review of access to services for people from particular vulnerable groups

Component	Description
Review question	In adults (18 years and older) at risk of depression or (anxiety disorders) from particular vulnerable groups (older people, BME groups and men) do service developments and interventions which are specifically designed to promote access, increase the proportion of people from the target group who access treatment, when compared with standard care? (RQ3.0)
Population	Adults (18 years and older) identified as at risk of depression (or anxiety disorders*) from the following vulnerable groups older adults BME groups men *Note: due to limited depression specific evidence, a broader evidence base (including anxiety disorders) will be used. An update of the review conducted for the Common mental health problems: identification and pathways to care guideline (NICE 2011) will be undertaken
Intervention(s)	Service developments or changes which are specifically designed to promote access. Specific models of service delivery (that is, community-based outreach clinics, clinics or services in non-health settings). Methods designed to remove barriers to access (including stigma, misinformation or cultural beliefs about the nature of mental disorder)
Comparison	Standard care
Outcomes	Critical outcomes: proportion of people from the target group who access treatment uptake of treatment Important but not critical outcomes: satisfaction, preference

Update 2017

Component	Description
	anxiety about treatment
Study design	• Systematic reviews of RCTs • RCTs • Cluster RCTs

12.2.11 Clinical evidence

Potentially relevant papers were identified through several different sources including a specific search of the literature on modifications to facilitate access, a general search of the literature on psychological interventions for depression, a search of the literature published between 2009 and 2015 on interventions to treat depression, previous iterations of the guideline (NICE 2004, NICE 2009), existing systematic reviews (Beach, Gary et al. 2006, Kehle, Greer et al. 2011, Dorstyn, Saniotis et al. 2013) and handsearch.

Forty-nine papers (13 SRs and 36 RCTs) were reviewed at full text for this review. Of these 6 met eligibility criteria (including 3 SRs) (Beach, Gary et al. 2006, Kehle, Greer et al. 2011, Dorstyn, Saniotis et al. 2013) and led to the inclusion of 12 RCTs: (Callahan, Hendrie et al. 1994, Hedrick, Chaney et al. 2003, Oslin, Sayers et al. 2003, Bartels, Coakley et al. 2004, Beach, Gary et al. 2006, Dobscha, Corson et al. 2006, Ross, TenHave et al. 2008, Dwight-Johnson, Aisenberg et al. 2011, Kehle, Greer et al. 2011, Chong and Moreno 2012, Dorstyn, Saniotis et al. 2013, Naeem, Gul et al. 2015).

These RCTs cover strategies for the three special groups of interest (BME groups, men and older people) and represent each of the three types of intervention of interest (service developments, models of delivery and methods to remove barriers to access).

Further information about both included and excluded studies is contained within Appendix J10. The full GRADE evidence profiles and associated forest plots can be found in Appendices L and M.

Update 2017

12.2.11.1 Telephone-administered psychological interventions versus usual care

2 RCTs (Dwight-Johnson, Aisenberg et al. 2011, Chong and Moreno 2012) were identified that investigated the impact of telephone-administered psychological interventions compared with usual care, both of which were conducted in BME populations.

An overview of the trials included in the meta-analyses can be found in Table 242.

Summary of findings can be found in Table 243 and Table 244.

Data were available for all critical outcomes. No data were available for the important outcomes of preference and anxiety about treatment.

Table 242: Study information table for trials included in the meta-analysis of telephone administered psychological interventions versus usual care

	Clinic-based telepsychiatry using video-webcam versus usual care	Telephone CBT versus enhanced usual care
Total no. of studies (N[1])	1 (197)	1 (101)
Study ID	Chong 2012	Dwight-Johnson 2011
Country	USA	USA
Target group	BME (Hispanic)	BME (Hispanic)
Mean age in years (SD or range)	43 (11.9)	39.8 (10.6)
Disorder	Depression	Depression

	Clinic-based telepsychiatry using video-webcam versus usual care	Telephone CBT versus enhanced usual care
Gender (% male)	11	22
Intervention	Clinic-based telepsychiatry using an online virtual meeting programme (addressed following factors to target access: language and cultural concerns [Hispanic psychiatrists provided intervention]; cost [patients were not asked to pay for any MH services provided in the clinic])	Telephone CBT (CBT, translated into the Spanish language and checked for relevance to the local Latino context and culture): 8 sessions of 45-50mins
Comparison	TAU (care received from usual providers)	Enhanced usual care (any typically available care for depression, patients were encouraged to talk with their primary care provider about depression)
Notes: [1]N=total number of participants; TAU=treatment as usual; CBT=cognitive behavioural therapy		

Table 243: Summary of findings table for the comparison of clinic based telepsychiatry using a video webcam versus usual care for adults with depression from particular vulnerable groups (BME groups)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Clinic-based telepsychiatry using a video Webcam versus TAU (95% CI)
Number of subjects who made a mental health appointment Not reported	167 (1 study) 6 months	⊕⊖⊖⊖ very low[1,2,3] due to risk of bias, indirectness, imprecision	RR 2.89 (2.14 to 3.9)	**Study population**	
				333 per 1000	**630 more per 1000** (from 380 more to 967 more)
				Moderate	
				333 per 1000	**629 more per 1000** (from 380 more to 966 more)
Number of subjects who made a primary care appointment Not reported	167 (1 study) 6 months	⊕⊖⊖⊖ very low[1,2,3] due to risk of bias, indirectness, imprecision	RR 0.8 (0.68 to 0.94)	**Study population**	
				874 per 1000	**175 fewer per 1000** (from 52 fewer to 280 fewer)
				Moderate	
				874 per 1000	**175 fewer per 1000** (from 52 fewer to 280 fewer)
				Study population	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Clinic-based telepsychiatry using a video Webcam versus TAU (95% CI)
Number used antidepressants Not reported	167 (1 study) 6 months	⊕⊖⊖⊖ **very low**[1,2,3] due to risk of bias, indirectness, imprecision	**RR 1.52** (1.16 to 1.99)	**460 per 1000**	**239 more per 1000** (from 74 more to 455 more)
				Moderate	
				460 per 1000	**239 more per 1000** (from 74 more to 455 more)
Mean number of completed mental health appointments Not reported	106 (1 study) 6 months	⊕⊖⊖⊖ **very low**[1,2,4] due to risk of bias, indirectness, imprecision			The mean number of completed mental health appointments in the intervention groups was **0.5 higher** (0.94 lower to 1.94 higher)
Mean number of completed primary care appointments Not reported	132 (1 study) 6 months	⊕⊖⊖⊖ **very low**[1,2,5] due to risk of bias, indirectness, imprecision			The mean number of completed primary care appointments in the intervention groups was **0 higher** (1.17 lower to 1.17 higher)
Satisfaction Visit Specific Satisfaction Questionnaire (VSQ-9). Scale from: 0 to 36.	167 (1 study) 6 months	⊕⊖⊖⊖ **very low**[2,5,6] due to risk of bias, indirectness, imprecision			The mean satisfaction in the intervention groups was **0.2 higher** (0.16 lower to 0.56 higher)

[1] Unclear blinding of outcome assessment
[2] US study with potential applicability issues
[3] Events<300
[4] 95% CI crosses both line of no effect and threshold for clinically significant benefit (SMD 0.5)
[5] N<400
[6] Non-blind outcome assessment (self-report)

Table 244: Summary of findings table for the comparison of telephone CBT versus enhanced usual care for adults with depression from particular vulnerable groups (older people, BME groups and men)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Enhanced usual care	Risk difference with Telephone CBT (95% CI)
Number reporting they were satisfied with the treatment provided	97 (1 study)	⊕⊖⊖⊖ very low[1,2] due to risk of bias, imprecision	RR 1.03 (0.59 to 1.79)	**364 per 1000**	**11 more per 1000** (from 149 fewer to 287 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Enhanced usual care	Risk difference with Telephone CBT (95% CI)

Notes:
[1] High ROB in one domain and unclear ROB in two others
[2] 95% CI crosses two clinical decision thresholds

1

2

12.2.1.23 Telephone-administered monitoring interventions versus usual care

4 2 RCTs (Oslin, Sayers et al. 2003, Ross, TenHave et al. 2008) were identified that
5 investigated the impact of telephone-administered monitoring interventions compared with
6 usual care, both conducted in male populations, one of which was also in older adults.

7 An overview of the trials included in the meta-analyses can be found in Table 245.

8 Summary of findings can be found in Table 246 and Table 247.

9 Data were available for all critical outcomes. No data were available for the important
10 outcomes of satisfaction, preference and anxiety about treatment.

11 **Table 245: Study information table for trials included in the meta-analysis of**
12 **telephone-administered monitoring interventions versus usual care**

	Telephone disease management versus usual care	Close monitoring versus usual care
Total no. of studies (N[1])	1 (97)	1 (233)
Study ID	Oslin 2003	Ross 2008
Country	USA	USA
Target group	Older people/men	Men
Mean age in years (SD or range)	61.6 (10.5)	59.2 (15.9)
Disorder	Depression	Minor depression
Gender (% male)	96	93
Intervention	Telephone disease management programme (A behavioural health specialist [nurse] maintained regularly scheduled telephone contact to: develop a treatment plan; monitor treatment effectiveness and adverse effects; assess and encourage treatment adherence; offer support and education)	Close monitoring (telephone calls from health technician to: monitor symptoms of depression with PHQ-9; ask participants if they were currently interested in receiving treatment for their depressive symptoms)
Comparison	Usual care (including: education for providers on existing treatment guidelines; screening patients attending clinic; providing diagnostic information to the clinician; making general suggestions for treatment including encouraging clinicians to	Usual care (primary care clinicians were given a full report of the baseline assessment with suggestions for ongoing monitoring of depressive symptoms and had the option to request referral of patients to a mental health clinic; each patient also received a letter

	Telephone disease management versus usual care	Close monitoring versus usual care
	refer patients to the behavioural health clinic)	following their initial assessment that included self-help advice for any significant depression symptoms and encouragement to discuss symptoms with their primary care clinician

Notes:
[1] N=total number of participants

Table 246: Summary of findings table for the comparison of telephone disease management versus usual care for adults with depression from particular vulnerable groups (older people and men)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Telephone disease management versus usual care (95% CI)
Number completing at least one mental health/substance abuse appointment Self-report	97 (1 study) 4 months	⨁⊖⊖⊖ **very low**[1,2,3] due to risk of bias, indirectness, imprecision	RR 4.21 (1.71 to 10.37)	**Study population**	
				98 per 1000	**315 more per 1000** (from 70 more to 919 more)
				Moderate	
				98 per 1000	**315 more per 1000** (from 70 more to 918 more)

Notes:
[1] Non-blind outcome assessment (self-report)
[2] US study with potential applicability issues and veteran population so may not be applicable to all men
[3] Events<300

Table 247: Summary of findings table for the comparison of close monitoring versus usual care for adults with depression from particular vulnerable groups (men)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Close monitoring versus usual care (95% CI)
Number attending primary care visits during study period Case review	223 (1 study) 6 months	⨁⊖⊖⊖ **very low**[1,2,3] due to risk of bias, indirectness, imprecision	RR 1.06 (0.89 to 1.27)	**Study population**	
				667 per 1000	**40 more per 1000** (from 73 fewer to 180 more)
				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Close monitoring versus usual care (95% CI)
				667 per 1000	**40 more per 1000** (from 73 fewer to 180 more)
Number who had any MH care (including behavioural health specialist) during the study period Case review	223 (1 study) 6 months	⨁⊖⊖⊖ **very low**[1,2,4] due to risk of bias, indirectness, imprecision	RR 5.13 (2.28 to 11.54)	**Study population**	
				65 per 1000	**266 more per 1000** (from 83 more to 680 more)
				Moderate	
				65 per 1000	**268 more per 1000** (from 83 more to 685 more)
Number who started an antidepressant during the study period Case review	223 (1 study) 6 months	⨁⊖⊖⊖ **very low**[1,2,3] due to risk of bias, indirectness, imprecision	**RR 1.67** (0.8 to 3.48)	**Study population**	
				97 per 1000	**65 more per 1000** (from 19 fewer to 240 more)
				Moderate	
				97 per 1000	**65 more per 1000** (from 19 fewer to 241 more)

Notes:
[1] Outcome assessment was non-blind and there were statistically significant baseline differences between groups (more males, more financial troubles, more subjects with trauma exposure, more with a past history of depression and more with a GAD diagnosis in the intervention group)
[2] US study with potential applicability issues and veteran population so may not be applicable to all men
[3] 95% CI crosses both line of no effect and threshold for clinically significant benefit (RR 1.25)
[4] Events<300

Update 2017

12.2.1.31 Simple collaborative care versus usual care

3 RCTs (Callahan, Hendrie et al. 1994, Hedrick, Chaney et al. 2003, Dobscha, Corson et al. 2006) were identified that investigated the impact of simple collaborative care compared with usual care. Two of these RCTs were conducted in male populations and one in an older adult population.

An overview of the trials included in the meta-analyses can be found in Table 248.

Summary of findings can be found in Table 249.

Data were available for all critical outcomes. No data were available for the important outcomes of satisfaction, preference and anxiety about treatment.

Table 248: Study information table for trials included in the meta-analysis of simple collaborative care versus usual care

	Simple collaborative care versus usual care	
	Men	**Older adults**
Total no. of studies (N[1])	2 (729)	1 (175)
Study ID	Dobscha 2006[2] Hedrick 2003[3]	Callahan 1994
Country	USA	USA
Mean age in years (SD or range)	56.8 (11.0)2 57.2 (13.9)3	65.1
Disorder	Depression	Depression
Gender (% male)	93[2] 95[3]	76
Intervention	Depression decision support team (1 psychiatrist + 1 nurse care manager) provided 1 early patient educational contact and depression monitoring with feedback to clinicians[2] Mental health team provided a treatment plan to the primary care provider, telephoned patients to support adherence to the plan, reviewed treatment results, and suggested modifications to the provider)[3]	Specialist advice (3 additional GP visits, with instructions on referral and suggested clinical actions including suggestions about providing basic psychoeducation to the patient in the intervention letter from the study team)
Comparison	Usual care[2] Consultant liaison care[3]	TAU

Notes:
[1]N=total number of participants; TAU=treatment as usual
[2]Dobscha 2006
[3]Hedrick 2003

Table 249: Summary of findings table for the comparison of simple collaborative care versus usual care for adults with depression from particular vulnerable groups (older people and men)

Outcomes	**No of Participants (studies) Follow up**	**Quality of the evidence (GRADE)**	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with Control	**Risk difference with Simple collaborative care versus usual care (95% CI)**
Number who attended ≥1 appointment with mental health specialist Database review	729 (2 studies) 12 months	⨁⊖⊖⊖ **very low**[1,2,3,4] due to risk of bias, inconsistency, indirectness, imprecision	RR 1.2 (0.77 to 1.86)	**Study population**	
				323 per 1000	**65 more per 1000** (from 74 fewer to 277 more)
				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Simple collaborative care versus usual care (95% CI)
				323 per 1000	**65 more per 1000** (from 74 fewer to 278 more)
Number who have had a depression-related primary care visit Database review	354 (1 study) 12 months	⊕⊖⊖⊖ **very low**[1,3,5] due to risk of bias, indirectness, imprecision	RR 1.47 (1.28 to 1.7)	**Study population**	
				570 per 1000	**268 more per 1000** (from 160 more to 399 more)
				Moderate	
				570 per 1000	**268 more per 1000** (from 160 more to 399 more)
Number of patients whose unhelpful medications (those potentially exacerbating depression) were terminated	175 (1 study)	⊕⊖⊖⊖ **very low**[6,7] due to risk of bias, imprecision	**RR 1.01** (0.58 to 1.76)	**227 per 1000**	**2 more per 1000** (from 95 fewer to 172 more)
Received ≥ 90 days of therapy with a minimally therapeutic dosage of antidepressant Database review	625 (2 studies) 12 months	⊕⊖⊖⊖ **very low**[1,2,3,4] due to risk of bias, inconsistency, indirectness, imprecision	**RR 1.13** (0.95 to 1.35)	**Study population**	
				605 per 1000	**79 more per 1000** (from 30 fewer to 212 more)
				Moderate	
				610 per 1000	**79 more per 1000** (from 31 fewer to 214 more)
Number of adults starting an antidepressant	175 (1 study)	⊕⊕⊖⊖ **low**[5,6] due to risk of bias, imprecision	**RR 3.25** (1.41 to 7.5)	**80 per 1000**	**180 more per 1000** (from 33 more to 520 more)
Number of patients for whom a psychiatric consultation was sought	175 (1 study)	⊕⊖⊖⊖ **very low**[6,7] due to risk of bias, imprecision	**RR 0.82** (0.38 to 1.75)	**147 per 1000**	**26 fewer per 1000** (from 91 fewer to 110 more)

Notes:
[1] Statistically significant group differences at baseline in Hedrick 2003 (more subjects with previous depression in intervention group)
[2] I-squared > 50%
[3] US study with potential applicability issues and veteran population so may not be applicable to all men
[4] 95% CI crosses both line of no effect and threshold for clinically significant benefit (RR 1.25)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Simple collaborative care versus usual care (95% CI)
[5] Events<300 [6] Unclear ROB in multiple domains [7] 95% CI crosses two clinical decision thresholds					

12.2.1.41 Co-located versus geographically separate services

2 1 RCT (Bartels, Coakley et al. 2004) was identified that investigated the impact of co-locating
3 services rather than keeping them geographically separate (usual care). This RCT was
4 conducted in an older adult population.

5 An overview of the trials included in the meta-analyses can be found in Table 250.

6 Summary of findings can be found in Table 251.

7 Data were available for all critical outcomes. No data were available for the important
8 outcomes of satisfaction, preference and anxiety about treatment.

9 **Table 250: Study information table for trials included in the meta-analysis of co-**
10 **located versus geographically separate services**

	Co-located services versus geographically separate services (usual care)
Total no. of studies (N[1])	1 (2,022)
Study ID	Bartels 2004
Country	USA
Target group	Older adults
Mean age in years (SD or range)	73.5 (6.2)
Disorder	Depression
Gender (% male)	74
Intervention	Integrated care model: 1) mental health and substance abuse services co-located in the primary care setting (including assessment, care planning, counselling, case management, psychotherapy, and pharmacological treatment), with no distinction in terms of signage or clinic names; 2) specialist services provided by licensed providers; 3) communication about the clinical evaluation and treatment plan between the specialist and primary care provider; and 4) an appointment with the specialist provider within 2 to 4 weeks following the primary care visit.
Comparison	Enhanced referral model (referral within 2 to 4 weeks of the primary care provider appointment; 2) treatment offered in a separate location by licensed professionals; 3) coordinated follow-up contacts if the patient failed to make the first scheduled visit; 4) assistance with transportation; and 5) visit costs covered
Notes: [1]N=total number of participants	

Table 251: Summary of findings table for the comparison of co-located services versus geographically separate services for adults with depression from particular vulnerable groups (older people)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Geographically separate services	Risk difference with Co-located services (95% CI)
Number of patient who engaged with treatment	1297 (1 study)	⊕⊕⊕⊖ **moderate**[1] due to risk of bias	RR 1.46 (1.34 to 1.59)	**514 per 1000**	**237 more per 1000** (from 175 more to 304 more)
Number of treatment visits	1390 (1 study)	⊕⊕⊖⊖ **low**[1,2] due to risk of bias, imprecision		The mean number of treatment visits in the control groups was **2.22**	The mean number of treatment visits in the intervention groups was **1.28 higher** (0.87 to 1.69 higher)

Notes:
[1] Unclear ROB in multiple domains
[2] 95% CI crosses one clinical decision threshold

12.2.1.54 Culturally-adapted psychological interventions versus usual care

1 RCT (Naeem, Gul et al. 2015) was identified that investigated the impact of tailoring a psychological intervention to the culture of their target group as opposed to providing usual care. This RCT was conducted in a BME population.

An overview of the trials included in the meta-analyses can be found in Table 252.

Summary of findings can be found in Table 253.

No data were available for the critical outcomes of proportion of people accessing treatment or uptake of treatment, or the important outcomes of preference and anxiety about treatment.

Table 252: Study information table for trials included in the meta-analysis of culturally-adapted psychological interventions versus usual care

	Culturally-adapted CBT versus usual care
Total no. of studies (N[1])	1 (137)
Study ID	Naeem 2015
Country	Pakistan
Target group	BME
Mean age in years (SD or range)	31.7 (11.1)
Disorder	Depression
Gender (% male)	40
Intervention	Culturally adapted CBT (adjustments included a family member accompanying the participant, the addition of a family session, initial focus on physical symptoms, Urdu translations of jargon, culturally appropriate homework assignments and use of folk stories and examples relevant to local religious beliefs): 6 individual sessions plus 2 family sessions
Comparison	TAU (treatment as usual; typically medication and hospital visits)
Notes:	

	Culturally-adapted CBT versus usual care
[1]N=total number of participants	

Table 253: Summary of findings table for the comparison of culturally-adapted CBT versus TAU for adults with depression from particular vulnerable groups (BME groups)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Culturally-adapted CBT (95% CI)
Number of participants stating that they were 'very satisfied' with treatment	137 (1 study)	⊕⊖⊖⊖ **very low**[1,2] due to risk of bias, imprecision	RR 1.54 (1.15 to 2.06)	**471 per 1000**	**254 more per 1000** (from 71 more to 499 more)

Notes:
[1] High ROB in multiple domains
[2] 95% CI crosses one clinical decision threshold

12.2.24 Economic evidence

No economic evidence on service developments and interventions that have been specifically designed to promote access to services for vulnerable groups of adults with, or at risk of, depression was identified by the systematic search of the literature. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

12.2.30 Clinical evidence statements

12.2.3.11 Telephone administered psychological interventions versus usual care

- Very low quality evidence from 2 RCTs (k=1-1, n=97-167) showed that BME patients receiving clinic-based telepsychiatry using a video webcam were more likely to use antidepressants and made more mental health appointments than those receiving care as usual, and that there was a clinically important but not statistically significant increase in completed mental health appointments in the telepsychiatry group, but no difference in the number of primary care appointments either made or completed between the two groups. Additionally there was no difference in the level of satisfaction reported, including on the visit specific satisfaction questionnaire (VSQ-9), between those receiving telephone administered psychological interventions compared with usual care.

12.2.3.21 Telephone-administered monitoring interventions versus usual care

- Very low quality evidence from 1 RCT (k=1, n=97) found that more older men receiving telephone disease management completed at least one mental health or substance misuse appointment compared with those receiving usual care
- Very low quality evidence from 1 RCT (k=1, n=223) found that more men in the close monitoring group received mental health care (including appointments with behavioural specialists) compared with those receiving usual care; there was a clinically important but not statistically significant increase in the number of men commencing antidepressant treatment during the study period with close monitoring and there was no difference in the number of men attending a primary care visit for case review compared with usual care.

Update 2017

12.2.3.3 Simple collaborative care versus usual care

- Very low quality evidence from up to 2 RCTs (k=1-2, n=354-729) found no difference in the number of men who attended at least one mental health appointment or who received more than 90 days of an antidepressant at a minimally therapeutic dose in the simple collaborative care and usual care treatment groups, but that more men in the collaborative care group received a depression-related primary care visit than those receiving usual care.
- Very low-low quality evidence from 1 RCT (k=1, n=175) found no difference in the number of older people who had their potentially unhelpful medications (those that may be exacerbating depression) terminated or for whom a psychiatric consultation was sought between the collaborative care and usual care conditions, but that more older people in the collaborative care than the usual care group started an antidepressant.

12.2.3.4 Co-located versus geographically separate services

- Low-moderate quality evidence from 1 RCT (k=1-1, n=1,297-1,390) found greater numbers of older people engaged with treatment and a small increase in the number of treatment visits when they attended co-located rather than geographically separate services.

12.2.3.5 Culturally-adapted psychological interventions versus care as usual

- Very low quality evidence from 1 RCT (k=1, n=137) found that more patients reported that they were 'very satisfied' with treatment when they had received culturally-adapted interventions compared with care as usual.

12.2.4 Economic evidence statements

- No evidence on the cost effectiveness of service developments and interventions that have been specifically designed to promote access to services for vulnerable groups of adults with, or at risk of, depression is available.

Update 2017

12.3 From evidence to recommendations

12.3.1 Relative values of different outcomes

The GC identified the proportion of people from the target group who access treatment and take up treatment and improvements in depression symptomology, response, remission, relapse and acceptability (loss to follow-up) as the critical outcomes for this question. Satisfaction, preference and anxiety about treatment were identified as important outcomes.

No evidence for either of the important, but not critical, outcomes of anxiety about treatment or patient preference was found.

12.3.2 Trade-off between clinical benefits and harms

The GC noted that there is evidence from a secondary analysis of the Coventry 2014 review of collaborative care, that interventions delivered via telephone were as effective as those delivered face-to-face. They noted that this evidence was from the general mental health population rather than only from the specific groups of interest for this review question, but agreed that it would be appropriate to extrapolate from that evidence base. They also discussed the issue of patient choice, and the fact that some people (particularly older people) may not be comfortable using technology and may prefer a face-to-face intervention. They therefore recommended interventions be available in a range of different methods.

They discussed the fact that there is currently a drive within the NHS to provide services outside of standard working hours and that although evidence on uptake of this was mixed,

and cost-effectiveness has not been established, that practitioners have found evening appointments to be popular with patients.

The GC noted that a number of the interventions reviewed may have clinical benefits both directly, in terms of increased uptake of treatment, and indirectly in terms of greater satisfaction leading to better engagement with services. The GC noted that co-locating services with physical health services (in particular for older people), active monitoring (for men) and involving families (for BME patients) appeared to improve access to and uptake of services.

No evidence of harm related to any of the interventions reviewed was found but it is possible that co-location and more active or assertive monitoring may be experienced by some people as stigmatising and improved access could lead to more 'false positive' and unnecessary and burdensome assessments or interventions for some people. The GC did note that participants provided with a number of the reviewed interventions made more appointments (showing greater uptake) but did not necessarily keep these (suggesting poor engagement).

The GC also recognised, drawing on their own knowledge and experience and the successes of the national roll out of the Improving Access to Psychological Therapies programme that the development of robust systems for the delivery of care are associated with improved uptake of services. This is particularly the case when supported by clear protocols for assessment, supporting service user choice, self-referral, entry criteria, information sharing, care coordination and outcome monitoring. The GC noted that such systems commonly referred to as 'stepped care models' would promote effective integration of interventions in primary and secondary care for the treatment of people with more and less severe depression and therefore developed recommendations that specified what the care pathways should include and achieve.

Update 2017

12.3.3 Trade-off between net health benefits and resource use

No evidence on the cost-effectiveness of service developments and interventions that have been specifically designed to promote access to services for vulnerable groups of adults with, or at risk of, depression was identified and no further economic analysis was undertaken.

The GC acknowledged that enhanced accessibility to services and integrated delivery of services for people with depression across primary and secondary care are likely to have considerable resource implications. The GC noted, however, that facilitating timely access to effective and cost-effective NICE-recommended treatments for depression results in more efficient use of resources and better outcomes for service users; moreover, there may be significant cost-savings for the NHS and social care as delayed or poorly co-ordinated treatment may negate the need for more costly intensive treatments for entrenched or chronic depressive symptoms. The GC noted that availability of services out of normal hours (evenings/weekends) is already established and would not entail significant resource implications.

The GC also acknowledged that routine collection of data on access to, uptake of, and outcomes of the interventions in the pathway is likely to have moderate resource implications. However, they expressed the opinion that routine collection of such data will allow more effective planning, delivery and evaluation of services, leading to more efficient use of resources and enhanced equality within and across services.

12.3.4 Quality of evidence

The evidence for this review generally came from single studies, of low to very low quality, of a reasonable sample size. The evidence was generally direct, from patients with symptoms of depression. However a number of the studies were conducted in the USA where

healthcare is structured very differently and there are additional issues relating to accessing services, such as financial considerations and greater geographical distance. The evidence relating to telephone disease management in particular came from a single US study in a war-veteran population, and so may have limited applicability to a UK setting. These issues were considered by the GC when interpreting the evidence.

In the context of the limited evidence base the GC chose to make a recommendation for research into interventions that could increase engagement with services in groups who are under-represented in services treating people with depression (for example, in BME and LGBT groups, men and older people).

12.3.50 Other considerations

The GC were aware, based on their clinical experience and knowledge that there are certain vulnerable groups (such as older people, men and people from BME communities) who are less likely to access services for depression. The GC discussed whether it was possible to make recommendations tailored specifically to each of these groups of people that would improve their access to services for depression. However, given the limited evidence available, the GC did not think it was possible to do so. Instead the GC made general recommendations on what should be done to promote access and increased uptake of services, highlighting older people, men and people from LGBT people and BME communities as particular groups to be aware of.

The GC noted, despite concerns about depression and suicide in younger men, that no evidence was found for interventions to increase access for this particular group. In the absence of evidence about what may be effective for this group the committee were wary of making specific recommendations for practice using consensus. They agreed, however, that the recommendations made should improve access for younger men too.

In light of the limited evidence and concerns raised by GC members and stakeholders in the consultation on the scope for the guideline, the GC decided to make a research recommendation on what are the most effective and cost effective methods to promote increased access to, and uptake of, interventions for people with depression who are under-represented in current services.

Update 2017

12.4.10 Recommendations

122. Commissioners and providers of mental health services should consider using stepped care models for organising the delivery of care and treatment of individuals with depression. Stepped care pathways should:

- provide accessible information about the pathway, for example in different languages and formats
- are accessible and acceptable to people using the services
- support the integrated delivery of services across primary and secondary care
- have clear criteria for entry to the service
- have multiple entry points and ways to access the service, including self-referral
- have agreed protocols for sharing information. [new 2017]

123. Commissioners and providers of mental health services should ensure pathways are in place to support the coordination of care and treatment of individuals with depression. Pathways should:

- promote easy access to, and uptake of, interventions in the pathway

- allow for prompt assessment of adults with depression, including assessment of severity and risk
- provide access to NICE-recommended interventions for depression
- ensure coordination and continuity of care
- have routine collection of data on access to, uptake of, and outcomes of the interventions in the pathway. [new 2017]

124. Commissioners and providers of mental health services should ensure pathways have the following in place for people with depression (in particular for men, older people, lesbian, gay, bisexual and transgender people and people from black, Asian and minority ethnic communities) to promote access and increased uptake of services:

- information about the pathway provided in a non-stigmatising way, using age and culturally appropriate language and formats
- services available outside normal working hours
- a range of different methods to engage with and deliver interventions, for example text messages, email, telephone and online
- services provided in community-based settings, for example, in an individual's home, community centres, leisure centres, care homes, social centres and integrated clinics within primary care
- bilingual therapists or independent translators
- involvement of families/partners. [new 2017]

12.5 Research recommendations

7. What are the most effective and cost effective methods to promote increased access to, and uptake of, interventions for people with depression who are under-represented in current services?

Statement: A series of randomised controlled trials should be conducted to determine what are the most effective and cost effective methods for promoting access or treatment for people with depression. The studies should address the needs of groups who are under-represented in services including older people and people from black, Asian and minority ethnic communities.

Rationale: There is general under-recognition of depression but the problem is more marked in certain populations. In addition, even where depression is recognised by the person with depression or by health professionals, access to treatment can still be difficult. A number of factors may relate to this limited access including a person's view of their problems, the information available on services and the location, design and systems for referral to services. A number of studies have addressed this issue and a number of strategies have been developed to address it but no consistent picture has emerged from the research which can inform the design and delivery of services to promote access. Little is also known about how these systems might be tailored to the needs of particular groups such as older people, people from black, Asian and minority ethnic communities, and people with disabilities who may have additional difficulties in accessing services.

Update 2017

13 Economic modelling: cost effectiveness of interventions for relapse prevention

13.1 Introduction – objective of economic modelling

The choice of long-term maintenance therapy in people with depression that is in remission was identified by the GC and the guideline health economist as an area with potentially major resource implications. Existing economic evidence in this area was limited and did not cover all relevant interventions. The clinical evidence in the area of relapse prevention was judged to be sufficient and of adequate quality to inform primary economic modelling. Based on the above considerations, an economic model was developed to assess the relative cost effectiveness of interventions aiming at preventing relapses in adults with depression that is in remission in the UK.

It is noted that the term 'relapse' is typically used to refer to a new episode of depression following incomplete or only brief recovery (e.g. less than 4 months of being well), whereas the term 'recurrence' usually means a new episode following a period of recovery lasting more than 4 months. Also, 'remission' is defined as a relatively brief period during which an improvement of sufficient magnitude is observed so that the individual no longer meets syndromal criteria for the disorder and has no more than minimal symptoms, whereas 'recovery' is defined as an extended asymptomatic phase, which lasts more than 6 months. In this chapter, the term 'relapse' is used to capture new depressive episodes occurring either within or beyond 4 months of an asymptomatic (recovery) phase and the terms 'remission' and 'recovery' are used interchangeably to capture any period where a person with depression no longer meets syndromal criteria for the disorder, regardless of the duration of this period.

Update 2017

13.2 Methods

13.2.1 Population

The study population of the economic model comprised adults with depression that is in full remission, following treatment for an acute depressive episode. People with partial remission or residual symptoms were not included in the analysis, as they constitute a distinct group for which evidence in the area of relapse prevention is rather limited.

The economic analysis focused on populations treated in primary care, as this is the setting where the majority of the study population is treated in routine practice. Moreover, populations treated in secondary care may have more severe and complex depression including comorbidities, so some aspects of care may be more difficult to determine and quantify in economic modelling. On the other hand, the GC acknowledged that the vast majority of RCTs in the area of relapse prevention have been conducted in secondary care settings. This may suggest that the study populations had a higher level of severity of depression, or may simply reflect clinical practice patterns at the time and in the countries in which the RCTs were conducted. Due to lack of relevant data from primary care settings, efficacy data were derived from RCTs conducted in secondary care and this is acknowledged as a limitation of the data and the economic analysis.

The GC suggested that the economic model take account of different predictors of relapse in depression, such as age, severity of initial depression, residual symptoms, psychiatric comorbidities, and number of previous episodes. However, identifying different sub-groups according to predictors of relapse within the evidence base was beyond the scope of the review question on relapse prevention.

Nevertheless, the number of previous depressive episodes is a well-established predictor of relapse (Keller & Shapiro, 1981; Kessing & Andersen, 1999; Mueller et al., 1999; Solomon et al., 2000) and therefore this factor was explored further in the context of the economic analysis. The majority of RCTs included in the guideline systematic review of interventions for relapse prevention provided some information on the minimum or mean number of previous episodes experienced by the study participants, and these details were used to identify studies in people with low risk of relapse (no previous depressive episodes), medium risk of relapse (1-2 previous episodes) and high risk of relapse (3+ previous episodes), as suggested by the GC (Table 254). Very few studies included participants who had remitted from their first depressive episode. Some studies provided information on interventions tested in participants with a mean of 1-2 previous episodes. The majority of trials included participants with a mean number of episodes that was greater than 3. Some studies did not provide any information on the number of previous episodes experienced by the study participants. These data were too sparse to indicate a differential treatment effect according to the number of previous episodes. However, since the number of previous episodes is a predictor of relapse, the economic analysis considered populations with a medium risk of relapse (1-2 previous episodes) and a high risk of relapse (3+ previous episodes) to explore the impact of relapse preventive interventions on costs and benefits according to the number of previous episodes experienced by the study population. The number of previous episodes experienced by each population determined their baseline risk of relapse (i.e. the risk of relapse under standard care and without the assessed intervention) and also the range of interventions assessed in the economic model, as determined by available evidence (for example, some interventions, such as mindfulness-based cognitive therapy (MBCT), have been tested only in populations with a high risk of relapse, as determined by a number of at least 3 previous episodes). Due to sparseness of relevant data, the same treatment effect was used in the two populations (that is, at medium and high risk of relapse, respectively, according to their number of previous depressive episodes).

In order to quantify epidemiological parameters and estimate economic model inputs, the base-case analysis for people with 1-2 previous episodes utilised baseline relapse data for people with 1 previous episode, and the analysis for people with 3+ episodes utilised baseline relapse data on people with 3 previous episodes.

Regarding the severity of the depressive episodes, the economic analysis assumed that people at medium risk of relapse would experience less severe depression if they relapsed and populations at high risk of relapse would experience more severe depression if they relapsed. The definition of less severe and more severe depression was used to classify the study populations in the review questions on interventions for the treatment of a new episode of depression and is provided in section 7.2. This distinguishing of populations in this economic analysis was reflected only in the utility values of the remission state considered in the economic model structure, owing to lack of efficacy data specific to symptom severity level. People with less severe depression were assumed to always experience less severe depression if they relapsed over the duration of the analysis; similarly, populations with more severe depression were assumed to always experience more severe depression if they relapsed over the time horizon of the model. This assumption was necessary in order to populate the economic model. The selection of populations in terms of risk and severity of depression aimed to cover a wide range of adults with depression that is in remission presenting in routine clinical practice.

Based on the above categorisations of the study population, the following scenarios were tested in economic analysis for people treated in primary care:

- People at medium risk of relapse (1-2 previous episodes) who were assumed to experience less severe depression if they relapsed
- People at high risk of relapse (3+ previous episodes) who were assumed to experience more severe depression if they relapsed

Update 2017

In a scenario explored in sensitivity analysis, people at medium risk of relapse were assumed to experience more severe depression if they relapsed, and people at high risk of relapse were assumed to experience less severe depression if they relapsed.

The cohorts assessed in the economic model were divided into sub-groups, depending on the acute treatment they had received for their depressive episode that led to remission of the episode. Three broad cohort categories were assessed, reflecting the availability of clinical data: cohorts that achieved remission following acute pharmacological treatment with antidepressants; cohorts that achieved remission following acute psychological treatment; and cohorts that remitted following acute combined psychological and pharmacological treatment. People that had achieved remission following antidepressant drug treatment were further sub-divided into 4 sub-groups according to the class/type of antidepressant they had been receiving as acute treatment: SSRI, SNRI, TCA and mirtazapine, respectively. Cohorts that had remitted following less commonly used antidepressants (e.g. nefazodone, maprotiline, mianserine, phenelzine or reboxetine) or other treatments such as lithium or ECT and cohorts that remitted from treatment-resistant depression were not assessed in the economic analysis, due to sparseness of relevant data and the fact that these sub-groups represent a smaller part of the study population (so they were considered as of lower priority for economic analysis).

Table 254: Population characteristics in relapse prevention RCTs included in the guideline systematic review and considered in the economic analysis

Study ID	Comparison	Number of previous episodes		Risk of relapse
		Inclusion criterion?	Mean (SD)	
SSRIs received prior to randomisation				
Wilson 2003	Sertraline vs placebo	Not relevant	1st for 72.5%	Low
Keller 1998		chronic major or double	1.8 (2.2)	Medium
Doogan1992		Not relevant	79% ≥ 1 previous episode	Medium or high
Lepine 2004		min 2 episodes in last 4 years	50% ≥ 6 episodes	High
McGrath 2006	Fluoxetine vs placebo	Not relevant	Not reported	?
Gilaberte 2001		min 1 episode in last 5 years	2.4 (1.2) - 1.1 in last 5years	Medium
Reimherr 1998		Note relevant	median 1-1.5	Medium
Schmidt 2000		72% had previous episodes	Not reported	Medium or high
Montgomery 1988		min 1 episode in last 5 years	2.3 in last 5 years, 3.8 total	High
Petersen 2010		chronic (≥36months) or min 3 episodes or history of poor recovery or double depression	4.2 (5.6)	High
Perlis 2002	Fluoxetine vs fluoxetine + CT	chronic (≥36months) or min 3 episodes or history of poor recovery or double depression	4.4 (5.9); 5.6 (9.2)	High
Terra 1998	Fluvoxamine vs placebo	min 2 episodes in last 5 years	3.5 (1.4)	High
Gorwood 2007	Escitalopram vs placebo	Not relevant	Not reported	?
Rapaport 2004		Not relevant	Not reported	?
Kornstein 2006		min 2 episodes, 1 in last 5 years	4.7 (3.1); 5.8 (6.0)	High
Montgomery 1992	Citalopram vs placebo	Not relevant	Not reported	?
Robert 1995		Not relevant	Not reported	?
Klysner 2002		Not relevant	85% 1st episode; max 2 previous	Low
Hochstrasser 2001		min 2 episodes, 1 in last 5 years	4 (2-15); 3 (2-20)	High
Hollon 2005	Paroxetine (± lithium or desipramine) vs placebo	Not relevant	Not reported	?
Montgomery 1993		min 2 episodes in last 4 years	20% 2; 56% 3-4; 24% 5+	High
SNRIs received prior to randomisation				

Update 2017

Study ID	Comparison	Number of previous episodes		Risk of relapse
		Inclusion criterion?	Mean (SD)	
Perahia 2006	Duloxetine vs placebo	min 1 episode	Not reported	Medium or high
Perahia 2009		min 2 episodes in last 5 years	4 (1.5); 4.4 (2.3)	High
Simon 2004	Venlafaxine vs placebo	Not relevant	Not reported	?
Montgomery 2004		min 1 episode in last 5 years	1.4 (0.7) in past 5 years	Medium
PREVENT study		min 2 episodes, 1 in last 5 years	Not reported	Likely high
PREVENT studyb		min 2 episodes, 1 in last 5 years	Not reported	Likely high
TCAs received prior to randomisation				
Coppen 1978a	Amitriptyline vs placebo	Not relevant	34% 1st, 66% 2nd or 3rd	Medium
Alexopoulos 2000	Nortriptyline vs placebo	Not relevant	30% 1st, 47.5% 2nd, 14% 3rd	Medium
Georgotas 1989		Not relevant	3.9 (5.3); 3.5 (6.5)	High
Stewart 1997	Imipramine (± lithium) vs placebo	chronic depression (min 2 years)	Not reported	?
Prien 1984		Not relevant	median 4	High
Mirtazapine received prior to randomisation				
Thase 2001	Mirtazapine vs placebo	min 1 episode in past 5 years or chronic depression	recurrent 54%	Medium or high
Any AD received prior to randomisation				
Segal 2010	MBCT + AD taper vs AD	min 2 previous episodes	4.7 (2.3)	High
Kuyken 2008		min 2 previous episodes	median 6; 35% ≥ 9	High
Kuyken 2015		min 2 previous episodes	46% ≥ 5	High
Huijbers 2015	MBCT + AD vs AD	Not relevant	mean 7.4; median 5	High
Huijbers 2016	MBCT + AD vs MBCT + AD taper	min 2 previous episodes	5.9 (5.3) - 5.6 (4.1)	High
Wilkinson 2009	group CBT + AD vs AD	Not relevant	31% 1st, 20% 2nd, 31% 3-5, 18% >5	Medium
CT received prior to randomisation				
Jarrett 2001	CT vs no treatment	min 1 previous episode	2.3 (0.15)	Medium
Jarrett 2013	CT vs fluoxetine vs placebo	min 1 previous episode	median 3	High
CBASP received prior to randomisation				

Update 2017

Study ID	Comparison	Number of previous episodes		Risk of relapse
		Inclusion criterion?	Mean (SD)	
Klein 2004	CBASP vs assessment only	chronic (≥2 years) or recurrent with incomplete remission between episodes or + dysthymic disorder or chronic + dysthymic disorder	2.4 (1.6)	Medium
IPT or IPT plus AD received prior to randomisation				
Frank 2007	Weekly vs biweekly vs monthly IPT	recurrent depression	mean 5	High
Combination therapy received prior to randomisation				
Reynolds 2006	IPT + paroxetine vs IPT vs paroxetine vs placebo	Not relevant	55% first episode	Low
Petersen 2010	CBT + fluoxetine vs CBT + placebo	chronic (≥ 36mths) or min 3 episodes or history of poor recovery or double depression	2.3 (1.5); 8.6 (15.1)	High
Frank 1990	IPT + imipramine vs IPT vs imipramine vs placebo	min 2 episodes	mean 6.8 (7.3), median 4	High
Reynolds 1999	IPT + nortriptyline vs IPT vs nortriptilyne vs placebo	min 1 episode in past 3 years	median 4	High
TAU received prior to randomisation				
Bockting 2005	group CT + TAU vs TAU	min 2 episodes in last 5 years	88%+75%>2; median 3.5	High
Godfrin 2010	MBCT + TAU vs TAU [Williams 2014 included a 3rd arm of attention control]	min 2 episodes	Not reported	Likely high
Bondolfi 2010		min 2 episodes in past 5 years, 1 in past 2 years	median 4	High
Ma 2004		min 2 episodes in past 5 years, 1 in past 2 years	median 3	High
Meadows 2014		min 2 episodes (10% BD)	8.1 (7.7); 11.4 (16.4); median 5	High
Teasdale 2000		min 2 episodes in past 5 years, 1 in past 2 years	median 3	High
Williams 2014		min 2 episodes in past 5 years, 1 in past 2 years	77% >4 previous	High
Stangier 2013	CBT + TAU vs psychoeducation + TAU	min 3 episodes	7.4 (8.3)	High

Notes:
Risk of relapse defined as follows: 1st episode suggests low risk; 1-2 previous episodes suggest medium risk; 3+ previous episodes suggest high risk
Interventions: AD: antidepressant; CBASP: cognitive behavioural analysis system of psychotherapy; CBT: cognitive behavioural therapy; CT; cognitive therapy; IPT: interpersonal psychotherapy; MBCT: mindfulness-based cognitive therapy; SNRI: serotonin-norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor; TAU: treatment as usual; TCA: tricyclic antidepressant

Update 2017

Study ID	Comparison	Number of previous episodes		Risk of relapse
		Inclusion criterion?	Mean (SD)	
Other abbreviations: Min: minimum; max: maximum; SD: standard deviation				
1				

Update 2017

Starting age of modelled population

The age of cohorts considered in the economic model was determined by the mean age of onset of depression in adults and the number of previous episodes that people experienced. Kessler et al. (2005) reported the results of a national comorbidity household survey in the US, according to which the median age-of-onset of depression was 32 years (interquartile range 19-44 years). In a Swedish longitudinal cohort study of 3,563 people followed up for 30-49 years, the median age at first onset of depression was reported to be around 35 years (Mattisson et al., 2007). A large (n=20,198) Scottish family-based population study designed to identify the genetic determinants of common diseases, including major depression disorder, reported a mean age of onset of major depressive disorder of 31.7 years (SD 12.3 years) among 2,726 participants that met DSM-IV criteria for current and/or past major depression disorder (Fernandez-Pujals et al., 2015). On the other hand, Andrade et al. (2003) did a review of results of community epidemiological surveys on major depressive episodes that were carried out in 10 countries in America, Europe and Asia (UK was not included in these countries); the authors reported a median age of onset of major depression in the early to mid-twenties in all countries other than Japan (late twenties) and the Czech Republic (early thirties). Based on this evidence and following GC expert advice, the age of onset of major depression in the cohorts considered in the model was set at 32 years.

According to the GC expert opinion, the mean interval between 2 consecutive depressive episodes in people who experience relapses is about 2 years. Therefore, for modelling purposes, people with 1 previous episode remitting from their current episode were assumed to be 34 years old, and people with 3 previous episodes remitting from their current episode were assumed to be 38 years of age.

Percentage of women in the study population

The percentage of women in each cohort were estimated to be 56%, based on weighted epidemiological data on depressive episodes reported in the most recent adult psychiatric morbidity household survey conducted in England (McManus et al., 2016).

Determining the age and gender mix of the cohorts was necessary in order to estimate mortality risks in the model.

Update 2017

13.2.2 Interventions assessed

The range of interventions assessed in the economic analysis was determined by the availability of relevant clinical data. Maintenance pharmacological treatments comprised commonly used antidepressants including SSRIs (citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine and sertraline), SNRIs (duloxetine and venlafaxine), TCAs (amitriptyline, nortriptyline and imipramine) and mirtazapine. Maintenance psychological treatments included MBCT, individual cognitive behavioural therapy (CBT) and individual or group cognitive therapy (CT). Combined psychological and pharmacological maintenance treatment was represented by individual CBT and fluoxetine; this was selected by the GC as the most representative and commonly used treatment in the NHS among combination treatments tested in the RCTs included in the guideline systematic review (other combination therapies tested in relapse prevention RCTs included IPT and paroxetine, IPT and imipramine, and also IPT and nortriptyline).

Comparators included no maintenance treatment (waitlist) and clinical management, which reflects placebo trial arms and comprises visits to health professionals without any active pharmacological or psychological intervention being received (but with possible antidepressant drug tapering, if an antidepressant had been received as acute treatment).

Different interventions were assessed in people who had received pharmacological, psychological, or combined treatment as acute therapy that led to remission, according to the availability of respective clinical data and their risk for future relapses.

People who had remitted following acute pharmacological treatment moved on to one of the following maintenance treatment options:

- Cohorts at medium risk of relapse (1 previous episode):
  - continuation of the same drug they had been receiving as acute treatment, i.e. an SSRI, SNRI, TCA, or mirtazapine. Each class was represented in the analysis by the most commonly used antidepressant within the class. For SSRIs this was citalopram; for SNRIs venlafaxine; and for TCAs amitriptyline (Prescribing & Medicines Team, 2016; unpublished CPRD data provided by GC).
  - gradual discontinuation of antidepressant treatment (tapering) and clinical management comprising general practitioner (GP) visits; this option reflected care in RCT placebo arms. It needs to be noted that discontinuation of antidepressant was done abruptly in the placebo arms of some RCTs that informed the economic analysis, i.e. placebo replaced the drug immediately, while in other studies the drug was tapered and eventually replaced by pill placebo. Antidepressants are associated with withdrawal symptoms if they are discontinued abruptly, thus increasing the relative effect of maintenance antidepressant treatment, meaning that the overall treatment effect of maintenance antidepressant treatment versus antidepressant tapering may have been exaggerated in the clinical review and, consequently, in the economic analysis.
- Cohorts at high risk of relapse (3 previous episodes):
  - continuation of the same drug they had been receiving as acute treatment; as data for this analyses were derived mostly from studies assessing a mixture of antidepressants (therefore no drug-specific efficacy data were available), the economic analysis used citalopram for costing purposes, because this is the most commonly used antidepressant for the treatment of depression in adults (Health and Social Care Information Centre, 2016).
  - gradual discontinuation of antidepressant treatment (tapering) and clinical management comprising GP visits
  - gradual discontinuation of antidepressant treatment (tapering) and initiation of MBCT
  - combination therapy comprising continuation of drug treatment and addition of MBCT
  - combination therapy comprising continuation of drug treatment and addition of group CT

The last 3 options were considered only in cohorts at high risk of relapse because they have been tested specifically in populations with a high number of previous depressive episodes, and thus at high risk of relapse, in the trials included in the guideline systematic review.

People who had received acute psychological treatment prior to remission, represented by CT, as this was the intervention for which most evidence was available in this cohort, moved on to one of the following maintenance treatment options:

- Cohorts at medium risk of relapse (1 previous episode):
  - maintenance psychological treatment with CT
  - maintenance pharmacological treatment, represented by fluoxetine, as this was the only drug for which evidence was available in this population
  - clinical management, comprising GP visits, reflected in RCT placebo arms
  - no treatment, reflecting RCT wait list arms
- Cohorts at high risk of relapse (3 previous episodes):
  - maintenance psychological treatment with CT
  - maintenance pharmacological treatment, represented by fluoxetine

Update 2017

- clinical management, comprising GP visits
- no treatment
- MBCT
- group CT

The last 2 options were considered only in cohorts at high risk of relapse because they have been tested specifically in populations with a high number of previous depressive episodes, and thus at high risk of relapse, in the trials included in the guideline systematic review. Combination treatment was not assessed in people who had remitted following psychological acute treatment, due to lack of relevant evidence.

People who had received acute combination treatment prior to remission, represented by CBT and fluoxetine for the reasons discussed earlier, moved on to one of the following maintenance treatment options:

- Cohorts at high risk of relapse (3 previous episodes):
  - maintenance combination treatment, represented by individual CBT and fluoxetine
  - maintenance pharmacological treatment, represented by fluoxetine
  - gradual discontinuation of pharmacological treatment (tapering) and maintenance psychological treatment, represented by individual CBT
  - gradual discontinuation of antidepressant treatment (tapering) and clinical management, comprising GP visits, reflecting RCT placebo arms

All options were applied exclusively to cohorts at high risk of relapse, as defined by their number of previous episodes, because the largest part of this evidence came from populations with a high number (3+) of previous episodes.

It needs to be noted that a number of interventions included in the guideline systematic review of relapse prevention studies (shown in Table 254) have not been considered in the economic analysis. These include:

- maintenance group CBT combined with an antidepressant versus an antidepressant alone in people who had remitted following acute antidepressant treatment. Evidence from this intervention comes from a small pilot study conducted in older adults that was underpowered and showed insignificant effects (Wilkinson 2009, N=45). The GC expressed the opinion that this evidence was too thin to support any recommendations and therefore it was excluded from further consideration in the economic analysis.
- Maintenance CT plus fluoxetine versus fluoxetine alone in people who had received fluoxetine as acute treatment. Evidence came from a single, relatively small, RCT (Perlis 2002, N=132) that showed no significant effect of the combination therapy (RR 0.93, 95% CI 0.63 to 1.39). This suggests that the addition of CT on fluoxetine treatment adds extra cost for no benefit and therefore is not cost-effective. Thus this intervention and related evidence was also excluded from further consideration in the economic analysis.
- Maintenance CBASP versus assessment only in people who had received CBASP as acute treatment. Evidence came from a single, relatively small RCT (Klein 2004, N=82) judged to be of very low quality. This evidence was judged to be too thin to inform the economic modelling. Therefore this intervention and related evidence was excluded from further consideration.
- CBT plus treatment as usual versus psychoeducation plus treatment as usual, in people who were under treatment as usual at randomisation. Evidence came from a single RCT (Stangier 2013, N=180). Although this evidence was relevant, it was not possible to be incorporated into the economic analysis, because the interventions were not linked to the network of interventions in the network meta-analyses (NMAs) that were conducted to provide the economic model with efficacy data. Moreover, the study did not include a control intervention representing the baseline risk of relapse that would allow a separate

Update 2017

economic sub-group analysis informed by this trial. Therefore these interventions and related evidence were not considered further in the economic analysis.

- Weekly versus biweekly versus monthly IPT in people who received IPT or a combination of IPT with antidepressants at randomisation. Evidence came from a single RCT (Frank 2007, N=131). This evidence was not possible to incorporate into the economic analysis because the interventions were not linked to the network of interventions in the NMAs that were undertaken to inform the economic model. Moreover, the study did not include a control intervention representing the baseline risk of relapse that would allow a separate economic sub-group analysis informed by this trial. Therefore, assessment of different frequencies of IPT sessions was not carried out in the economic analysis.

13.2.31 Model structure

A Markov model was constructed using Microsoft Office Excel 2013. The model estimated the total costs and benefits associated with provision of each of the treatment options in each cohort of adults with depression that is in remission. The structure of the model, which aimed to simulate the course of depression and relevant clinical practice in the UK, was also driven by the availability of clinical data.

According to the model structure, hypothetical cohorts of adults with depression that is in full remission were initiated on relevant treatment options, according to the type of acute treatment they had received, as described in section 13.2.2. Separate models were developed for the various sub-populations considered in the analysis, depending on the type of the acute treatment of the depressive episode that led to remission of the episode.

The model, which was run in yearly cycles, included 3 health states: relapse (depressive episode), remission, and death. Within each year, people could remain in the same state or move from one state to another, with the exception of death, which was an absorbing state (so people in this state always remained in it). For every new episode of relapse, people entered separate relapse states (i.e. separate depressive episodes) so that their number of previous episodes could be tracked and the appropriate future risk of relapse that is dependent on the number of previous episodes could be applied. In addition, within each new episode of relapse, people entered tunnel relapse states, so that the time they remained in every relapse (depressive episode) could be estimated and a time-dependent probability of remission could be applied. People achieving remission also entered tunnel remission states, so that the time they remained in remission could be estimated and a time-dependent probability of relapse could be applied.

The time horizon of the analysis was 10 years, which allowed assessment of longer-term costs and benefits associated with relapse prevention treatment without introducing high complexity associated with the number of tunnel states that would be required were the model run over a longer period of time. A half-cycle correction was applied; this practically means that all events in the model occurred in the middle of each cycle.

Maintenance pharmacological (antidepressant) treatment was received during the first 2 years of the model; maintenance psychological treatment was received for the first year of the model. Cohorts under combined maintenance treatment received the pharmacological component of combined therapy during the first 2 years of the model and the psychological treatment component during the first year. Benefits of all treatments were assumed to be enjoyed over the first 2 years of the model, according to available evidence on pharmacological and psychological interventions aiming at relapse prevention and the GC expert opinion. Therefore, over the first 2 years in the model, the risk of relapse experienced by the cohorts was determined by their baseline risk of relapse and the efficacy of the maintenance treatment option received by each cohort. If people relapsed during this period of 2 years, maintenance treatment was discontinued and the preventative benefit of maintenance treatment ceased at the point of relapse. Beyond the period of the first 2 years, all cohorts were subject to the same baseline risk of relapse according to their number of

Update 2017

previous episodes and the time (years) spent in remission. The model did not assess future maintenance treatments beyond those received over the first 1-2 years of the model.

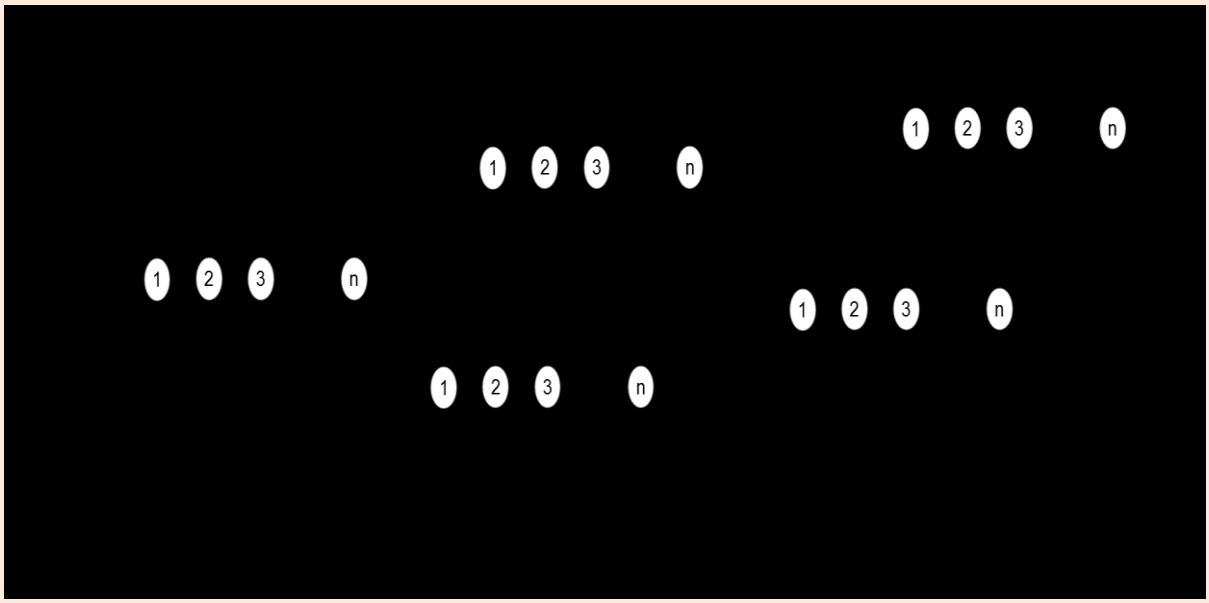
The baseline risk of relapse for each cohort depended on the time people remained in remission (the longer people stayed in remission, the lower their risk of relapse) and their number of previous episodes (the higher the number of their previous episodes, the higher their risk of relapse). The probability of remission for each cohort depended on the time people remained in relapse / a depressive episode (the longer people stayed in relapse, the lower their probability of remission).

The model did not consider probabilities and events associated with conversion to bipolar depression. This is a potential outcome that was not considered in the model due to sparseness of relevant data and the complexity entailed in modelling this outcome and associated future events.

People who received maintenance pharmacological treatment were assumed to experience common antidepressant side effects (such as headaches, nausea, agitation, sedation, or sexual dysfunction) resulting in a reduction in their HRQoL over the period of 2 years during which they received maintenance antidepressant treatment. They were also assumed to incur extra costs for the management of their side effects, which comprised GP visits and pharmacological treatment.

The structure of the economic model of relapse prevention is shown in Figure 21.

Figure 21. Schematic diagram of the relapse prevention economic model structure



Update 2017

13.2.4 Costs and outcomes considered in the analysis

The economic analysis adopted the perspective of the NHS and personal social services, as recommended by NICE (NICE, 2014). Costs consisted of intervention costs (drug acquisition, staff time for provision of maintenance pharmacological, psychological and combined therapies), as well as other costs associated with the management of future relapses, which included drug acquisition, primary care, hospitalisation, outpatient visits, psychological therapies, and accident and emergency visits. Costs of management of common side effects from antidepressants in people receiving maintenance pharmacological treatment alone or in combination and healthcare costs incurred by people in remission (potentially unrelated to the treatment of depression) were also considered in the analysis. The cost year was 2016.

The measure of outcome was the Quality Adjusted Life Year (QALY), which incorporated utilities associated with the health states of remission or relapse, as well as utility decrements due to common side effects associated with maintenance antidepressant treatment.

13.2.54 Efficacy data

13.2.5.15 Selection of efficacy data and methods of evidence synthesis

Efficacy data (expressed as numbers of people relapsing) for the relapse prevention interventions considered in the economic modelling were derived from the RCTs included in the respective guideline systematic reviews. As the study population in the economic models comprises adults with depression that is in full remission, the GC initially advised that only RCTs where participants were in full remission at randomisation be utilised in the model. A large proportion of RCTs included in the guideline systematic review used a more relaxed definition of remission as an inclusion criterion pre-randomisation, with a MADRS or HAMD cut-off point that was 2-3 points higher than the widely accepted thresholds for remission. Although the populations in these studies were not in full remission according to a stricter definition of remission, the GC accepted that this increase in the threshold for remission might not be clinically significant and also did not affect substantially the relative effect of treatment in these populations, as confirmed by inspection of the results in studies with a 'strict' versus those with a 'looser' definition of remission. Therefore the GC decided to include these studies in the economic analysis, in order to enhance the evidence base and help populate different branches of the economic models. Since this criterion was relaxed, a few trials that selected people who had responded to treatment at randomisation, some of whom were likely remitters, were also included in the analysis. Studies that included a mixture of people in full or partial remission were also included in the meta-analyses that informed the economic model. However, RCTs where all participants had residual symptoms were excluded from the economic analysis. Studies on older adults were not excluded from the economic analysis, in line with their inclusion in the clinical analysis of RCT data.

Update 2017

Drug-specific efficacy data inputs for the economic analysis of people at medium risk of relapse that had remitted following acute pharmacological treatment were obtained from pairwise meta-analysis of respective clinical data; details are provided in section 13.2.5.2. For all other analyses, data were synthesised in NMAs conducted within a Bayesian framework using Markov Chain Monte Carlo simulation techniques implemented in WinBUGS 1.4.3 (Lunn et al., 2000; Spiegelhalter et al., 2003). A binomial likelihood and cloglog link linear model was used (Dias et al., 2011) to allow estimation of hazard ratios of each maintenance treatment versus placebo, which were then applied onto the baseline risk of relapse in the first and second year of the economic analyses (after this period people returned to the baseline risk of relapse that corresponded to their number of previous episodes and the number of years spent in remission). Although, as discussed in Section 13.2.6, the risk of relapse in people with depression that is in remission is reduced over time following a Weibull distribution, the cloglog link linear model was appropriate to use; this is because hazard ratios between interventions are assumed to be constant over time, the shape parameter gamma of the Weibull distribution does not vary with time and, also, because in each RCT considered in the NMA, events across arms referred to the same follow-up time point.

The WinBUGS code used to synthesise the data, for both random and fixed effect models, is shown in Table 255. It is a simplified code compared with the 'standard' cloglog link linear model (Dias et al., 2011) in that the time parameter has been removed since hazard ratios are time-independent and events in each study refer to the same follow-up time. Depending on data availability, in each NMA fixed and/or random effect models were tested, as appropriate. Goodness of fit of each model was tested using the total residual deviance (totresdev) and the deviance information criteria (DIC) tool. Details on the interventions, data

and type of model used (i.e. fixed or random effects) in each NMA are reported in the respective sections 13.2.5.3, 13.2.5.4 and 13.2.5.5.

Table 255. WinBUGS codes used to synthesise data in all NMAs that informed the guideline economic modelling of interventions aiming at preventing relapses in people with depression that is in remission

Binomial likelihood, cloglog link

Random Effects model

```
# Binomial likelihood, cloglog link
# Random effects model for multi-arm trials
model{                          # *** PROGRAM STARTS
for(i in 1:ns){                 # LOOP THROUGH STUDIES
   w[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
   delta[i,1] <- 0            # treatment effect is zero for control arm
   mu[i] ~ dnorm(0,.0001)          # vague priors for all trial baselines
   for (k in 1:na[i]) {            # LOOP THROUGH ARMS
      r[i,k] ~ dbin(p[i,k],n[i,k]) # Binomial likelihood
# model for linear predictor
      cloglog(p[i,k]) <- mu[i] + delta[i,k]
      rhat[i,k] <- p[i,k] * n[i,k] # expected value of the numerators
#Deviance contribution
      dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
         +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))        }
#  summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
   for (k in 2:na[i]) {            # LOOP THROUGH ARMS
# trial-specific LHR distributions
      delta[i,k] ~ dnorm(md[i,k],taud[i,k])
# mean of LHR distributions, with multi-arm trial correction
      md[i,k] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
# precision of LHR distributions (with multi-arm trial correction)
      taud[i,k] <- tau *2*(k-1)/k
# adjustment, multi-arm RCTs
      w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
# cumulative adjustment for multi-arm trials
      sw[i,k] <- sum(w[i,1:k-1])/(k-1)
     }
  }
totresdev <- sum(resdev[])          #Total Residual Deviance
d[1]<-0      # treatment effect is zero for reference treatment
# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,0.01) }
sd ~ dunif(0,5)     # vague prior for between-trial SD
tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)
# pairwise HRs and LHRs for all possible pair-wise comparisons, if nt>2
for (c in 1:(nt-1)) {
for (k in (c+1):nt) {
lhr[c,k] <- (d[k]-d[c])
log(hr[c,k]) <- lhr[c,k]
}
}
} # *** PROGRAM ENDS
```

Update 2017

Binomial likelihood, cloglog link

Fixed Effects model

```
# Binomial likelihood, cloglog link
# Fixed effects model for multi-arm trials
model{                          # *** PROGRAM STARTS
for(i in 1:ns){                 # LOOP THROUGH STUDIES
   mu[i] ~ dnorm(0,.0001)          # vague priors for all trial baselines
   for (k in 1:na[i]) {            # LOOP THROUGH ARMS
      r[i,k] ~ dbin(p[i,k],n[i,k]) # Binomial likelihood
# model for linear predictor
      cloglog(p[i,k]) <- mu[i] + d[t[i,k]] - d[t[i,1]]
      rhat[i,k] <- p[i,k] * n[i,k]        # expected value of the numerators
#Deviance contribution
      dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
         +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))        }
#  summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
}
totresdev <- sum(resdev[])           #Total Residual Deviance
d[1]<-0      # treatment effect is zero for control arm
# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,0.01) }

# pairwise HRs and LHRs for all possible pair-wise comparisons, if nt>2
for (c in 1:(nt-1)) {
for (k in (c+1):nt) {
lhr[c,k] <- (d[k]-d[c])
log(hr[c,k]) <- lhr[c,k]
}
}
}              # *** PROGRAM ENDS
```

Each WinBUGS model was run with an initial burn-in period of 50,000 iterations, followed by 50,000 further iterations, thinned by 5 so as to obtain 10,000 iterations for use in the probabilistic economic model.

The models utilised uninformative prior parameters. Three different sets of initial values were used and convergence was tested by visual inspection of the Brooks Gelman-Rubin diagram. In addition, convergence of the models was assessed by checking the autocorrelation and the Kernel density plots within WinBUGS.

13.2.5.2 Efficacy data for people at medium risk of relapse who remitted following acute pharmacological treatment

Efficacy data for class-specific pharmacological treatments in people with depression at medium risk of relapse who remitted following acute pharmacological treatment were derived from placebo-controlled pharmacological relapse prevention RCTs in populations that had remitted following acute and/or continuation pharmacological treatment that were included in the guideline systematic review; it needs to be noted that some pharmacological relapse prevention studies randomised participants that were in remission after acute treatment and prior to continuation phase, whereas other studies had a different design and randomised participants that were in remission following a continuation phase and prior to a maintenance phase of treatment. The GC advised that continuation and maintenance phase studies be

Update 2017

analysed together. In all cases study endpoint data were used. Class treatment effects were used for SSRIs (represented by citalopram), SNRIs (represented by venlafaxine), and TCAs (represented by amitriptyline).

Endpoint treatment effects, in the form of risk ratios, as estimated in guideline pairwise meta-analysis, were applied onto the baseline relapse risk over the first 2 years of the economic analysis, during which pharmacological maintenance treatment was received. After the two years of maintenance pharmacological treatment people in the model returned to the baseline risk of relapse that corresponded to their number of previous episodes and the number of years they spent in remission.

Table 256 shows the RCTs, interventions and relative effects considered in the analysis of people at medium risk of relapse who remitted following acute pharmacological treatment, as well as the relative treatment effect (risk ratio) of each antidepressant class or mirtazapine versus placebo (which represented clinical management in the model), according to the guideline systematic review and meta-analysis in the area of pharmacological relapse prevention.

Table 256: RCTs, interventions and relative effects considered in the analysis of people at medium risk of relapse who remitted following acute pharmacological treatment

Intervention assessed in economic analysis	Intervention assessed in RCTs (all versus pill placebo)	Study IDs	Mean risk ratio (95% CIs)*
SSRIs (represented by citalopram)	Sertraline	Doogan 1992; Keller 1998; Lepine 2004; Wilson 2003	0.61 (0.56 to 0.68)
	Fluoxetine	Gilaberte 2001; McGrath 2006; Montgomery 1988; Petersen 2010; Reimherr 1998; Schmidt 2000	
	Fluvoxamine	Terra 1998	
	Escitalopram	Gorwood 2007; Kornstein 2006; Rapaport 2004	
	Citalopram	Klysner 2002; Hochstrasser 2001; Montgomery 1992; Robert 1995	
	Paroxetine (± lithium / desipramine)	Hollon 2005; Montgomery 1993	
SNRIs (represented by venlafaxine)	Duloxetine	Perahia 2006; Perahia 2009	0.66 (0.55 to 0.78)
	Venlafaxine	Montgomery 2004; PREVENT study; PREVENT studyb; Simon 2004	
TCAs (represented by amitriptyline)	Amitriptyline	Coppen 1978a	0.70 (0.43 to 1.14)
	Nortriptyline	Alexopoulos 2000; Georgotas 1989	
	Imipramine ± lithium	Prien 1984; Stewart 1997	
Mirtazapine	Mirtazapine	Thase 2001	0.67 (0.45 to 0.98)
Overall antidepressant effect			0.63 (0.58 to 0.68)

*as estimated in guideline pairwise meta-analysis

Update 2017

13.2.5.31 Efficacy data for people at high risk of relapse who remitted following acute pharmacological treatment

Efficacy data for people with depression at high risk of relapse who remitted following acute pharmacological treatment were derived from synthesis of data obtained from psychological and pharmacological relapse prevention RCTs in populations that had remitted following acute and/or continuation pharmacological treatment that were included in the guideline systematic review.

Psychological RCTs in these populations assessed maintenance psychological interventions instead of, or in addition to, antidepressants; these studies did not use specific antidepressant drugs (or classes), so that no class-specific effect could be obtained for antidepressants. In order to synthesise psychological and pharmacological study data, an overall antidepressant treatment effect of the 4 drug classes (SSRIs, SNRIs, TCAs and mirtazapine) was estimated out of all studies (pharmacological and psychological) and utilised in the analysis. This overall treatment effect was applied to citalopram, which was the drug used in this analysis in terms of acquisition cost. It is noted that inspection of antidepressant class-specific efficacy data suggests that the treatment effect is broadly similar across antidepressant drug classes (Table 256), so use of an overall antidepressant effect appeared to be reasonable.

In addition to the above studies, a number of studies considered maintenance psychological treatments in people under treatment as usual (as seen in Table 254), which comprised a range of treatments that could include no treatment, help from the family doctor or other routine healthcare if requested, antidepressant use, or depression relapse active monitoring. In order to incorporate this evidence into the economic analysis, these studies were included in the data synthesis for people at high risk of relapse who remitted following acute pharmacological treatment in a sensitivity analysis. As in this population treatment as usual comprises antidepressant treatment, the relative effect of psychological intervention plus treatment as usual versus treatment as usual alone that was estimated in these studies was assumed to equal the relative effect of the psychological intervention plus antidepressant versus antidepressant alone.

Update 2017

Data from the above studies were synthesised in two NMAs (one for the base-case analysis and one for the sensitivity analysis that included additional comparisons) using the cloglog link linear model, as described earlier. Both random and fixed effects models were tested. It should be noted that the efficacy data included in the NMA reflected intervention and study endpoints for some RCTs; for other RCTs, these data reflected study endpoints that were beyond intervention endpoints, as some studies did not report end of intervention data.

Studies, interventions and efficacy data included in the guideline systematic review that were considered in the NMA of interventions for people at high risk of relapse who remitted following acute pharmacological treatment are shown in Table 257. The networks of interventions included in the NMAs, both in the base-case and sensitivity analysis, are shown in Figure 22.

Table 257: RCTs, interventions and efficacy data (number of relapses [n] and number randomised [N] in each arm) considered in the analysis of people at high risk of relapse who remitted following acute pharmacological treatment

Study ID	Comparison	Data time point (weeks)	Arm 1		Arm 2		Arm 3	
			n	N	n	N	n	N
Doogan1992	Sertraline (arm 1) vs placebo (arm 2)[1]	44	77	185	74	110	NA	NA
Keller 1998		74	42	77	60	84	NA	NA
Lepine 2004		78	81	196	53	103	NA	NA
Wilson 2003		100	25	56	31	57	NA	NA

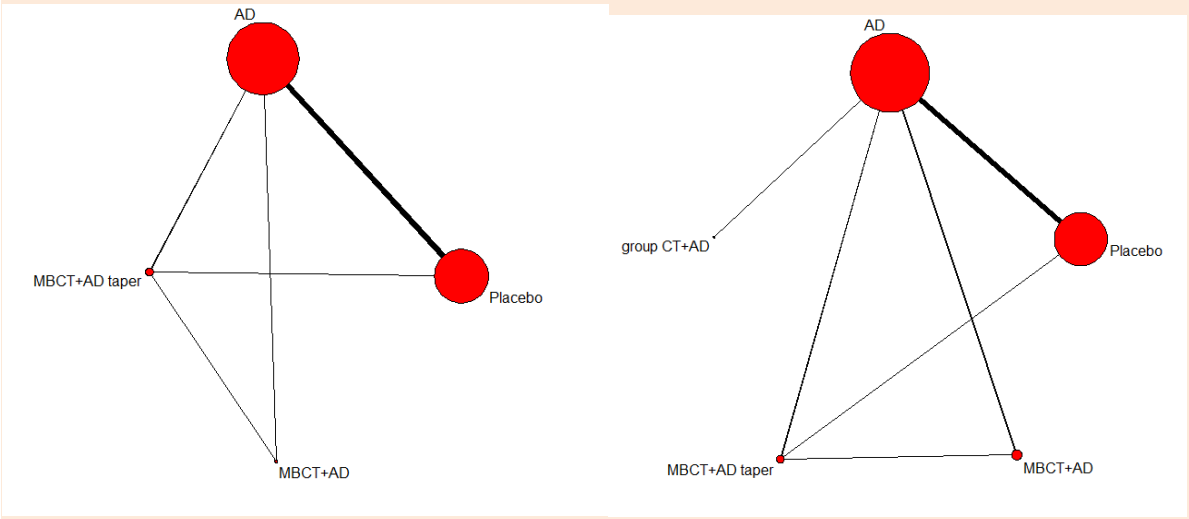
Study ID	Comparison	Data time point (weeks)	Arm 1		Arm 2		Arm 3	
			n	N	n	N	n	N
Gilaberte 2001	Fluoxetine (arm 1) vs placebo (arm 2)[1]	48	21	70	41	70	NA	NA
McGrath 2006		26	46	131	81	131	NA	NA
Montgomery 1988		52	43	108	72	112	NA	NA
Petersen 2010		80	5	15	9	17	NA	NA
Reimherr 1998		12	79	299	47	96	NA	NA
Schmidt 2000		25	105	189	87	122	NA	NA
Terra 1998	Fluvoxamine (arm 1) vs placebo (arm 2)[1]	52	14	110	33	94	NA	NA
Gorwood 2007	Escitalopram (arm 1) vs placebo (arm 2)[1]	24	23	152	63	153	NA	NA
Kornstein 2006		52	20	73	43	66	NA	NA
Rapaport 2004		36	89	181	62	93	NA	NA
Klysner 2002	Citalopram (arm 1) vs placebo (arm 2)[1]	48	37	60	55	61	NA	NA
Hochstrasser 2001		48	31	132	65	137	NA	NA
Montgomery 1992		24	30	105	23	42	NA	NA
Robert 1995		24	21	152	18	74	NA	NA
Hollon 2005	Paroxetine (± lithium / desipramine) (arm 1) vs placebo (arm 2)[1]	52	16	34	27	35	NA	NA
Montgomery 1993		52	11	68	29	67	NA	NA
Perahia 2006	Duloxetine (arm 1) vs placebo (arm 2)[1]	26	62	136	95	142	NA	NA
Perahia 2009		52	50	146	69	142	NA	NA
Simon 2004	Venlafaxine (arm 1) vs placebo (arm 2)[1]	26	79	161	101	157	NA	NA
PREVENT study		52	98	164	135	172	NA	NA
PREVENT studyb		52	12	43	25	40	NA	NA
Montgomery2004		52	24	109	64	116	NA	NA
Coppen 1978a	Amitriptyline (arm 1) vs placebo (arm 2)[1]	52	3	16	5	16	NA	NA
Alexopoulos 2000	Nortriptyline (arm 1) vs placebo (arm 2)[1]	104	4	22	11	21	NA	NA
Georgotas 1989		52	10	13	16	23	NA	NA
Prien 1984	Imipramine (± lithium) (arm 1) vs placebo (arm 2)[1]	104	13	40	22	34	NA	NA
Stewart 1997		26	9	17	8	15	NA	NA
Thase 2001	Mirtazapine (arm 1) vs placebo (arm 2)[1]	40	25	77	41	84	NA	NA
Kuyken 2008	MBCT (AD taper) (arm 1) vs AD (arm 2)	64	31	61	40	62	NA	NA
Kuyken 2015		104	94	212	100	212	NA	NA
Segal 2010	MBCT (AD taper) (arm 1) vs AD (arm 2) vs placebo (arm 3)	78	15	26	20	28	24	30
Huijbers 2015	MBCT + AD (arm 1) vs AD (arm 2)	64	17	33	20	35	NA	NA
Huijbers 2016	MBCT + AD (arm 1) vs MBCT (AD taper) (arm 2)	64	85	121	105	128	NA	NA
Bockting 2005	*group CT + TAU (arm 1) vs TAU (arm 2)*[2]	52	43	97	49	90	NA	NA
Williams 2014[3]	*MBCT + TAU (arm 1) vs TAU (arm 2)*[2]	52	55	108	31	56	NA	NA
Godfrin 2010		56	24	52	39	54	NA	NA

Update 2017

Study ID	Comparison	Data time point (weeks)	Arm 1		Arm 2		Arm 3	
			n	N	n	N	n	N
Bondolfi 2010		60	13	31	11	29	NA	NA
Ma 2004		60	14	37	23	38	NA	NA
Meadows 2014		60	42	101	53	102	NA	NA
Teasdale 2000		60	43	76	52	69	NA	NA

Notes:
1 These comparisons were treated in the network meta-analysis as 'antidepressant versus placebo'
2 These comparisons (and respective trials) were utilised only in sensitivity analysis; their relative effect was assumed to reflect the relative effect of 'intervention plus antidepressant' versus 'antidepressant alone'
3 This study included a third arm of attention control, which was of no interest to decision-making and did not offer any additional information in the network in terms of indirect comparisons; therefore, the information from this arm was excluded from the NMA and economic analysis.

Figure 22. Network of interventions included in the NMA of treatments for people at high risk of relapse who remitted following acute pharmacological treatment – base-case (left) and sensitivity (right) analysis



Results of the network meta-analysis: people at high risk of relapse who remitted following acute pharmacological treatment

The random effects model demonstrated a better fit for the data, for both the base-case and the sensitivity analysis. For the base-case analysis, with 77 data points (study arms) included in the NMA, the random effects model showed a better fit (totresdev = 79.66; DIC = 483.58) compared with the fixed effects model (totresdev = 91.38; DIC = 485.23). Similarly, for the sensitivity analysis, with 91 data points (study arms) included in the NMA, the random effects model showed a better fit (totresdev = 94.13; DIC = 569.14) compared with the fixed effects model (totresdev = 105.20; DIC = 570.34).

The results of the random effects models that informed the economic analysis are shown in Table 258. The table includes also results from direct head-to-head comparisons in the trials that informed the NMA (last column), to allow comparisons between NMA results and direct evidence. Results between the NMA and head-to-head comparisons are not directly comparable, because the NMA output was in the form of hazard ratios and results of direct, pairwise meta-analysis are expressed as risk ratios; however, it can be seen that NMA and pairwise meta-analysis results are overall consistent in direction and uncertainty around the mean effects.

Update 2017

Table 258. Results of the NMA that informed the economic analysis for people at high risk of relapse who remitted following acute pharmacological treatment (random effects model)

Comparison	Mean hazard ratio (95% CrI) - NMA	Mean risk ratio (95% CI) - pairwise meta-analysis
Base-case analysis		
AD vs placebo	0.51 (0.46 to 0.56)	0.63 (0.58 to 0.69)
MBCT (AD taper) vs placebo	0.45 (0.33 to 0.59)	0.72 (0.50 to 1.05)
MBCT + AD vs placebo	0.35 (0.22 to 0.52)	Not available
MBCT (AD taper) vs AD	0.88 (0.65 to 1.16)	0.88 (0.75 to 1.03)
MBCT + AD vs AD	0.68 (0.43 to 1.03)	0.90 (0.58, 1.40)
MBCT + AD vs MBCT (AD taper)	0.78 (0.52 to 1.13)	0.86 (0.74 to 0.99)
Standard deviation (NMA): mean 0.14 (95% CrI 0.02 to 0.28)		
Total residual deviance (NMA): mean 79.66 (95% CrI 57.19 to 103.70)		
Sensitivity analysis		
AD vs placebo	0.51 (0.46 to 0.56)	0.63 (0.58 to 0.69)
MBCT (AD taper) vs placebo	0.45 (0.34 to 0.58)	0.72 (0.50 to 1.05)
MBCT + AD vs placebo	0.35 (0.27 to 0.43)	Not available
Group CT + AD vs placebo	0.39 (0.23 to 0.63)	Not available
MBCT (AD taper) vs AD	0.88 (0.68 to 1.12)	0.88 (0.75 to 1.03)
MBCT + AD vs AD	0.68 (0.55 to 0.83)	0.79 (0.69 to 0.89)
Group CT + AD vs AD	0.77 (0.45 to 1.23)	0.81 (0.61 to 1.09)
MBCT + AD vs MBCT (AD taper)	0.78 (0.59 to 1.02)	0.86 (0.74 to 0.99)
Group CT + AD vs MBCT (AD taper)	0.88 (0.49 to 1.49)	Not available
Group CT + AD vs MBCT + AD	1.15 (0.64 to 1.90)	Not available
Standard deviation (NMA): mean 0.12 (95% CrI 0.01 to 0.25)		
Total residual deviance (NMA): mean 94.13 (95% CrI 70.09 to 119.70)		

Update 2017

13.2.5.44 Efficacy data for people at medium or high risk of relapse who remitted following acute psychological treatment

Efficacy data for people at medium risk of relapse and people at high risk of relapse who had remitted following acute psychological treatment were derived from synthesis of data obtained from psychological relapse prevention RCTs in populations that had remitted following acute and/or continuation psychological treatment that were included in the guideline systematic review.

In addition, studies assessing maintenance psychological treatments in people under treatment as usual were also included in a sensitivity analysis. These studies (and interventions) were considered only in people at high risk of relapse, since they had been tested specifically in populations with at least 3 previous depressive episodes. As in this population treatment as usual comprises no treatment, the relative effect of psychological intervention plus treatment as usual versus treatment as usual alone that was estimated in these studies was assumed to equal the relative effect of psychological intervention versus no treatment.

Data from the above studies were synthesised in a NMA using the cloglog linear model, as already described. Due to the lack of mixed comparisons (i.e. lack of direct and indirect

evidence in the same comparison) in the network, a fixed effects model was used. A single NMA was run for both people at medium risk of relapse and those at high risk of relapse. Since the additional studies and comparisons introduced new interventions in the analysis and did not create any loops, one NMA was run for both base-case and sensitivity analysis (as the evidence considered in the sensitivity analysis did not affect the relative effects obtained from the base-case analysis). Efficacy data included in the NMA reflected study endpoints that were beyond intervention endpoints, as some studies did not report end of intervention data.

Studies, interventions and efficacy data included in the guideline systematic review that were considered in the NMA of interventions for people at medium or high risk of relapse who remitted following acute psychological treatment are shown in Table 259. The networks of interventions included in the NMAs, both in base-case and sensitivity analysis, are shown in Figure 23.

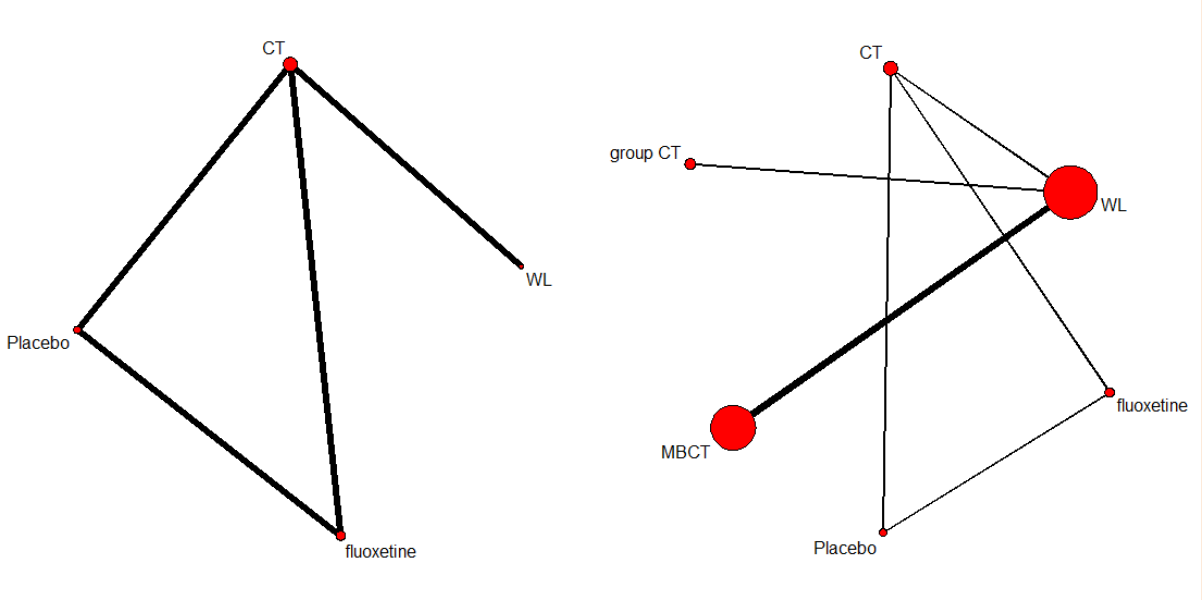
Table 259: RCTs, interventions and efficacy data (number of relapses [n] and number randomised [N] in each arm) considered in the analysis of people at medium and/or high risk of relapse who remitted following acute psychological treatment

Study ID	Comparison	Data time point (weeks)	Arm 1		Arm 2		Arm 3	
			n	N	n	N	n	N
Jarrett 2001	CT (arm 1) vs no treatment (arm 2)	56	14	41	22	43	NA	NA
Jarrett 2013	CT (arm 1) vs fluoxetine (arm 2) vs placebo (arm 3)	56	39	86	48	86	40	69
Bockting 2005	*group CT + TAU (arm 1) vs TAU (arm 2)*[1]	52	43	97	49	90	NA	NA
Williams 2014	*MBCT + TAU (arm 1) vs TAU (arm 2)*[1]	52	55	108	31	56	NA	NA
Godfrin 2010		56	24	52	39	54	NA	NA
Bondolfi 2010		60	13	31	11	29	NA	NA
Ma 2004		60	14	37	23	38	NA	NA
Meadows 2014		60	42	101	53	102	NA	NA
Teasdale 2000		60	43	76	52	69	NA	NA

1 These comparisons (and respective trials) were tested only in people at high risk of relapse, in a sensitivity analysis; their relative effect was assumed to reflect the relative effect of 'intervention' versus 'no treatment' (wait list)

Update 2017

Figure 23. Network of interventions included in the NMA of treatments for people at medium and/or high risk of relapse who remitted following acute psychological treatment – base-case (left) and sensitivity (right) analysis



Results of the network meta-analysis

The fixed effects model demonstrated a good fit for the data (totresdev = 19.81; DIC = 119.83, compared with 19 data points).

The results of the fixed effects model that informed the economic analysis are shown in Table 260. The table includes also results from direct head-to-head comparisons in the trials that informed the NMA (last column), to allow comparisons between NMA results and direct evidence. Results between the NMA and head-to-head comparisons are not directly comparable, because the NMA output was in the form of hazard ratios and results of direct, pairwise meta-analysis are expressed as risk ratios; however, it can be seen that NMA and pairwise meta-analysis results are overall consistent in direction and uncertainty around the mean effects.

Table 260. Results of the NMA that informed the economic analysis for people at medium risk of relapse and people at high risk of relapse who remitted following acute psychological treatment (fixed effects model)

Comparison	Mean hazard ratio (95% CrI) - NMA	Mean risk ratio (95% CI) - pairwise meta-analysis
CT vs placebo	0.72 (0.44 to 1.10)	0.78 (0.58 to 1.06)
Fluoxetine vs placebo	0.97 (0.61 to 1.47)	0.96 (0.73 to 1.27)
Wait list vs placebo	1.33 (0.54 to 2.82)	Not available
MBCT vs placebo [only sensitivity analysis]	0.91 (0.36 to 1.96)	Not available
Group CT vs placebo [only sensitivity analysis]	1.02 (0.36 to 2.32)	Not available
Fluoxetine vs CT	1.39 (0.88 to 2.08)	1.23 (0.91 to 1.66)
Waitlist vs CT	1.86 (0.89 to 3.57)	1.50 (0.89 to 2.51)
MBCT vs CT [only sensitivity analysis]	1.27 (0.59 to 2.48)	Not available
Group CT vs CT [only sensitivity analysis]	1.42 (0.59 to 2.98)	Not available
Wait list vs fluoxetine	1.41 (0.58 to 2.96)	Not available
MBCT vs fluoxetine [only sensitivity analysis]	0.96 (0.38 to 2.05)	Not available

Update 2017

Comparison	Mean hazard ratio (95% CrI) - NMA	Mean risk ratio (95% CI) - pairwise meta-analysis
Group CT vs fluoxetine [only sensitivity analysis]	1.07 (0.39 to 2.45)	Not available
MBCT vs wait list [only sensitivity analysis]	0.68 (0.55 to 0.83)	0.78 (0.68 to 0.89)
Group CT vs wait list [only sensitivity analysis]	0.76 (0.49 to 1.13)	0.81 (0.61 to 1.09)
Group CT vs MBCT [only sensitivity analysis]	1.13 (0.69 to 1.75)	Not available
Total residual deviance (NMA): mean 19.81 (95% CrI 11.37 to 31.92)		

13.2.5.5 Efficacy data for people at high risk of relapse who remitted following acute combined treatment

Efficacy data for people at high risk of relapse who remitted following acute combined psychological and pharmacological treatment were derived from synthesis of data obtained from RCTs in populations that had remitted following acute and/or continuation combined treatment that were included in the guideline systematic review. The studies for this population included in the review assessed a range of maintenance combined interventions (and their individual elements). Due to sparseness of data for specific interventions, the GC advised that relative effects of individual studies be combined and applied to any maintenance combination therapy and its components versus placebo. In the economic analysis, maintenance combined treatment (and its individual elements) for people remitting following acute combined treatment was represented by CBT and fluoxetine, as the most representative and commonly used treatment in the NHS among the combination treatments that were assessed in the RCTs included in the guideline systematic review.

Data from these RCTs were synthesised in a NMA, using the same cloglog linear model used in the other NMAs performed to inform the economic analysis of interventions for relapse prevention. A fixed effects model was used in this case, due to the small number of studies included in the analysis and the lack of mixed evidence in the network. In this set of studies interventions were provided for a long period of time (more than one year); studies reporting efficacy data over multiple time points indicated that the relative effect of maintenance treatment was higher at the end of first year and then was reduced over time; therefore, the NMA included efficacy data reported at study time points as close to 1 year as possible.

Studies, interventions and efficacy data included in the guideline systematic review that were considered in the NMA of maintenance interventions for people at high risk of relapse who remitted following acute combined treatment are shown in Table 261. The network of interventions included in the NMA is shown in Figure 24.

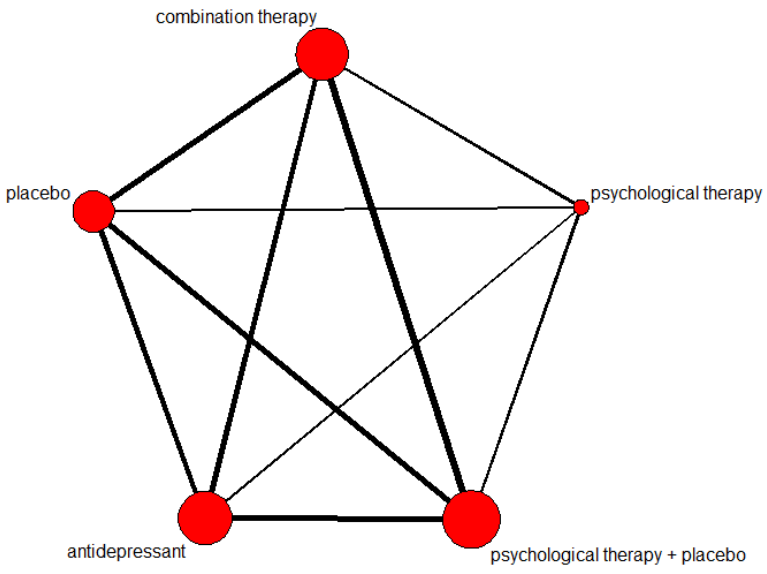
Table 261: RCTs, interventions and efficacy data (number of relapses [n] and number randomised [N] in each arm) considered in the analysis of people at high risk of relapse who remitted following acute combined treatment

Study ID	Comparison	Data time point (weeks)	Combined		Psych + placebo		Psych alone		Drug		Placebo	
			n	N	n	N	n	N	n	N	n	N
Frank 1990	IPT + imipramine vs IPT + placebo vs IPT vs imipramine vs placebo	52	4	25	14	26	14	26	11	28	18	23

Update 2017

Study ID	Comparison	Data time point (weeks)	Combined		Psych + placebo		Psych alone		Drug		Placebo	
			n	N	n	N	n	N	n	N	n	N
Reynolds 1999	IPT + nortriptyline vs IPT + placebo vs nortriptyline vs placebo	52	8	25	13	25	NA	NA	12	28	22	29
Petersen 2010	CBT + fluoxetine vs CBT + placebo	80	4	11	6	12	NA	NA	NA	NA	NA	NA
Reynolds 2006	IPT + paroxetine vs IPT + placebo vs paroxetine vs placebo	104	17	28	27	35	NA	NA	19	35	13	18

Figure 24. Network of interventions included in the NMA of treatments for people at high risk of relapse who remitted following acute combined treatment



Results of the network meta-analysis

The fixed effects model demonstrated a reasonable fit for the data (totresdev = 17.68; DIC = 78.15, compared with 15 data points).

The results of the fixed effects model that informed the economic analysis are shown in Table 262. The table includes also results from direct head-to-head comparisons in the trials that informed the NMA (last column), to allow comparisons between NMA results and direct evidence. Results between the NMA and head-to-head comparisons are not directly comparable, because the NMA output was in the form of hazard ratios and results of direct, pairwise meta-analysis are expressed as risk ratios; however, it can be seen that NMA and pairwise meta-analysis results are overall consistent in direction and uncertainty around the mean effects.

Update 2017

Table 262. Results of the NMA that informed the economic analysis for people at high risk of relapse who remitted following acute combined treatment (fixed effects model)

Comparison	Mean hazard ratio (95% CrI) - NMA	Mean risk ratio (95% CI) - pairwise meta-analysis
AD vs placebo	0.43 (0.27 to 0.64)	0.60 (0.46 to 0.78)
Psych therapy + placebo vs placebo	0.68 (0.44 to 1.00)	0.81 (0.60 to 1.11)
Psych therapy vs placebo	0.70 (0.34 to 1.27)	0.69 (0.45 to 1.04)
Combination therapy vs placebo	0.34 (0.20 to 0.52)	0.45 (0.19 to 1.04)
Psych therapy + placebo vs AD	1.64 (1.06 to 2.44)	1.35 (1.03 to 1.77)
Psych therapy vs AD	1.70 (0.80 to 3.11)	1.37 (0.77 to 2.45)
Combination therapy vs AD	0.81 (0.49 to 1.26)	0.82 (0.58 to 1.16)
Psych therapy vs psych therapy + placebo	1.06 (0.51 to 1.90)	1.00 (0.60 to 1.65)
Combination therapy vs psych therapy + placebo	0.51 (0.32 to 0.75)	0.65 (0.44 to 0.95)
Combination therapy vs psych therapy	0.53 (0.25 to 1.00)	0.30 (0.11 to 0.78)
Total residual deviance (NMA): mean 17.68 (95% CrI 11.82 to 27.38)		

13.2.64 Baseline risk of relapse

13.2.6.15 Baseline risk of relapse after a single (first) depressive episode (i.e. in people with no previous depressive episodes)

The baseline risk of relapse was estimated from data obtained from a review of long-term observational (or 'naturalistic' or 'longitudinal') studies conducted in primary or secondary care that reported relapse rates over long periods of time in people who had remitted from a depressive episode. In this type of studies the treatment is not assigned by design and is not under the control of the investigators. The review included 10 studies conducted in primary care (Coryell et al., 1991; Eaton et al., 2008; Hardeveld et al., 2013; Mattisson et al., 2007; Ormel et al., 1993; Riihimäki et al., 2014; Skodol et al., 2011; Stegenga et al., 2012; Van Weel-Baumgarten et al., 1998; Yiend et al., 2009) and 16 studies conducted in secondary care (Bukh et al., 2016; Gonzales et al., 1985; Holma et al., 2008; Kanai et al., 2003; Keller et al., 1984 and 1992; Keller and Shapiro,1981; Kennedy et al., 2003; Kiloh et al., 1988; Lee & Murray, 1988; Lehman et al., 1988; Maj et al., 1992; Melartin et al., 2004; Mueller et al., 1996 and 1999; Solomon et al., 2000) that reported relapse and/or chronicity data on people with depression. The studies were identified from 3 systematic reviews of naturalistic studies (Hardeveld et al., 2010; Steinert et al., 2014; Van Weel-Baumgarten et al., 2000) and further GC expert advice; additional studies were identified by scanning the reference lists of publications suggested by the GC.

The reported risks of relapse in the 1st year, 2nd to 5th years and 6th year and above following remission, together with risks of non-recovery over time reported in each study are provided in Table 263.

Update 2017

Table 263: Risks of relapse in years following remission and risks of chronicity of a depressive episode as reported in the naturalistic studies included in the guideline review

Study ID	Population characteristics	Relapse risk following remission			Chronicity (non-recovery)
		Year 1	Years 2-5	Years 6+	
Primary care – community settings					
Coryell et al., 1991	396 nonclinical individuals in the US who had had major depression that ended before the initial evaluation			Year 6: 0.34	
Eaton et al., 2008	92 adults with a first episode of major depression in a community setting in the US followed up for 10 years.	Graph: 0.06	Year 2: 0.25 (according to the graph, it is 0.19)	Year 10: 0.45	Year 10: 0.15 (chronicity defined as people not remaining free for longer than 1 year)
Hardeveld et al., 2013	687 people from the general Dutch population with a lifetime DSM-III-R diagnosis of major depression but without a current major depressive episode or dysthymia. Participants had to be at least 6 months in remission. 3-year follow-up & modelled projection of relapses	0.03	Year 2: 0.05 Year 5: 0.13	Year 10: 0.23 Year 20: 0.42	
Mattisson et al., 2007	Community sample of 3563 people in Sweden followed in 1947, 1957, 1972 & 1997. 344 people had their first onset of depression during the follow-up and were analysed in this study.	Graph: 0.09	Graph: Year 2: 0.12 Year 5: 0.21	Year 10: 0.29	
Ormel et al., 1993	20 people with depression among 201 people with common mental health problems receiving primary-care in the Netherlands				Year 3.5: 0.12
Riihimäki et al., 2014	137 people with DSM-IV depressive disorder in Finnish primary care; 122 completed a 5-year follow-up including 102 with a research diagnosis of major depression		Year 5: 0.51 [from full or partial remission]		Year 5: 0.10 (no full or partial remission) 0.31(no full remission)
Skodol et al., 2011	1,996 participants in a national US survey who met criteria for major depression, followed-up for 3 years	Not considered as only relapse after 1 year was estimated, those who relapsed in shorter periods of time were not included in estimates. Also, denominator included people with persistent major depression			Year 3: 0.15

Update 2017

Study ID	Population characteristics	Relapse risk following remission			Chronicity (non-recovery)
		Year 1	Years 2-5	Years 6+	
Stegenga et al., 2012	174 people with major depression in Dutch primary care, followed over 39 months.	0.11	Year 3: 0.18		Year 3: 0.17
Van Weel-Baumgarten et al., 1998	222 people with depression before January 1984 in Dutch primary care followed up for 10 years	Graph: 0.10	Graph: Year 2: 0.18 Year 3: 0.26 Year 5: 0.31	Year 10: 0.40	
Yiend et al., 2009	37 people attending UK primary care services followed for 23 years (73% with first episode); 23% on antidepressants at the time of the study (mean length of time on antidepressants during follow up 39.7 months); 24.3% received no pharmacological treatment. No patients were continuously medicated throughout follow up.			Year 10: 0.50 Year 23: 0.62	Year 23: 0.00
Secondary care – inpatient and/or outpatient settings					
Bukh et al., 2016	301 adult in- (60.8%) or out-patients with a validated diagnosis of a single depressive episode from 2005 to 2007 in Denmark	0.09	Year 2: 0.15 Year 5: 0.32		Year 1: 0.71 Year 2: 0.42 Year 5: 0.17
Gonzales et al., 1985	59 outpatients with unipolar major depression who had completed CBT and were followed for 1-3 years in the US	0.31			Year 1: 0.30
Holma et al., 2008	163 people in Finland with DSM-IV major depression receiving mainly outpatient care, followed up over 5 years between 1997 and 2004.		Year 5: 0.71		Year 5: 0.01 (no full or partial remission) 0.12 (no full remission)
Kanai et al., 2003	95 people who had recovered from unipolar major depression, followed for 6 years, recruited mostly from secondary settings (22/23 centres) in Japan. Participants had not received antidepressant or antipsychotic medication in the 3 months prior to the start of the study	0.21	Year 2: 0.30 Year 5: 0.42	Year 6: 0.14	
Keller & Shapiro, 1981	101 in- or out-patients in a current episode of major depression, of whom 75 recovered, followed for 1 year	0.21 (major depression)			Year 1: 0.26

Study ID	Population characteristics	Relapse risk following remission			Chronicity (non-recovery)
		Year 1	Years 2-5	Years 6+	
		0.36 (depressive symptoms)			
Keller et al., 1984	97 US people with an episode of major depressive disorder and no history of chronic minor depression who sought treatment at five university medical centres in the US				Year 2: 0.21
Kennedy et al., 2003	70 people receiving psychiatric secondary care, predominantly inpatient (76%) in the UK, with moderate to severe depression, followed up for 8-11 years. At follow up, 59% received at least 5 years of antidepressant treatment and only 15% received less than a year of antidepressant treatment. Over follow-up people maintained regular contact with their GPs and mental health teams for psychiatric review or treatment.	0.25	Year 2: 0.33	Graph: Year 8: 0.65	Year 11: 0.08
Kiloh et al., 1988	133 Australian inpatients with primary depressive illness between 1966 and 1970 were followed up for an average of 15 years.			Year 15: 0.76	Year 15: 0.17
Lee & Murray, 1988	89 inpatients with primary depressive illness in London in 1965-66 followed for 18 years			Year 18: 0.95	Year 18: 0.15
Lehman et al., 1988	65 depressed Canadians followed for 11 years; 52% were receiving psychiatric treatment predominately as outpatients at follow-up.			Year 11: 0.78	
Maj et al., 1992	72 people in specialist care in Italy who had recovered from an episode of non-psychotic major depression, evaluated bimonthly for a period ranging from 20 to 108 months (median 66 months).	0.37	Year 5: 0.75		
Melartin et al., 2004	269 secondary care psychiatric outpatients and inpatients diagnosed with a new episode of DSM-IV major depression in Finland		Year 1.5: 038		
Keller et al., 1992	431 people with major depression in secondary care in the US, followed for 10 years				Year 1: 0.33 Year 2: 0.19

Study ID	Population characteristics	Relapse risk following remission			Chronicity (non-recovery)
		Year 1	Years 2-5	Years 6+	
Mueller et al., 1996					Year 5: 0.12 Year 10: 0.07
Mueller et al., 1999	380 people who recovered from an index episode of major depressive disorder and 105 people who subsequently remained well for at least 5 years after recovery in outpatient specialist care in the US, followed for up to 15 years; people could be taking antidepressants and possibly ECT over time. Of those who eventually experienced a relapse, 77% were receiving no antidepressant treatment during the month just before the relapse.	Graph: 0.25	Graph: Year 2: 0.42 Year 3: 0.52	Year 15: 0.85 (Kaplan-Meier curve)	
Solomon 2000	318 people in inpatient and outpatient care in the US with unipolar major depressive disorder prospectively followed for 10 years Number of previous episodes: 0: 38%; 1: 24%; 2: 13%; 3+: 25% During the 4 weeks immediately before the onset of the first three prospectively observed relapses, 47%-50% of all subjects received no pharmacotherapy. During the 4 weeks immediately before the onset of the fourth and fifth prospectively observed relapses, one-third of the subjects received no pharmacotherapy.	0.25	Year 2: 0.42 Year 5: 0.60 2nd relapse: Year 2: 59% Year 5: 74% 3rd relapse: Year 2: 62% Year 5: 79% 4th relapse: Year 2: 62% 5th relapse: Year 2: 74% Number of relapses refer to prospectively observed relapses during the study, not lifetime relapses.		

Update 2017

1

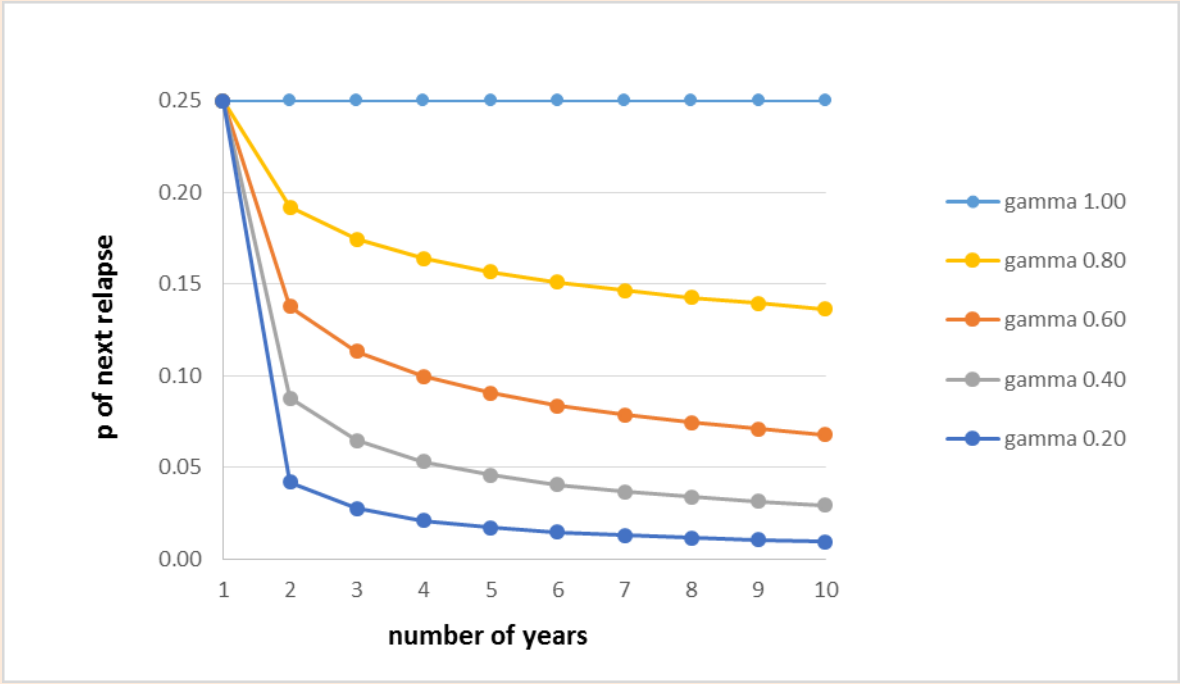
GC expert opinion and inspection of the available naturalistic data suggested that the risk of relapse of a depressive episode over time is dependent on time, and is likely to follow a Weibull distribution, in which the relapse rate is proportional to a power of time. People have a higher risk of relapse in the early years following remission, and this risk is reduced with every year they remain in remission; the cumulative hazard rate for the Weibull distribution is given by the following mathematical formula:

$$H(t) = \lambda t^{\gamma}$$

where lambda (λ) and gamma (γ) are the scale and shape parameters of the distribution, respectively.

When gamma >1, then the risk increases over time; when it equals 1, then the risk is constant with time and the distribution is exponential. When gamma < 1, then the risk is reduced over time. For example, the risk of relapse over time (years) from the previous depressive episode, for different rates of risk reduction (expressed by the gamma parameter) over time, assuming a first-year relapse risk of 0.25 (lambda = 0.25), is shown in Figure 25.

Figure 25. Change in risk of relapse over time from previous depressive episode, for different rates of risk reduction (expressed by a 'gamma' parameter) over time, and a first-year relapse risk of 0.25



Update 2017

Once people relapse and subsequently remit, their risk of relapse to the next episode increases again, and is dependent on the time they have spent in remission following resolution of their previous episode.

There is evidence that the risk of relapse increases with the number of previous episodes, and this was taken into account in the economic model (as described in section 13.2.6.2). Therefore, it was decided to estimate the baseline risk of relapse after the first depressive episode (i.e. in people with no previous depressive episodes) as a first step, and then model the baseline risk of relapse in the cohorts examined in the economic analysis according to their number of previous depressive episodes.

In order to estimate the risk of relapse over time and determine the underlying Weibull distribution after a single (first) depressive episode, the GC advised that data from Eaton et

al. (2008) and Mattisson et al. (2007) be synthesised; both studies included low-risk community cohorts, which were consistent with the model study population, who were followed up for long periods following remission of their first depressive episode. Both publications included graphs showing the time to relapse after the first episode of depression by gender. Digital software (http://www.digitizeit.de) was used to read and extract the proportions of people free from episode at each year of the study, up to 10 years. Subsequently, the numbers of people relapsing over time were approximated, based on the number of participants in each study. Data on men and women were similar, suggesting that there is no difference in the risk of relapse over time by gender. These data were synthesised in WinBUGS 1.4.3 (Lunn et al., 2000; Spiegelhalter et al., 2003) using a fixed effects model, in order to estimate the parameters of the underlying Weibull distribution (lambda and gamma). The model was run with an initial burn-in period of 20,000 iterations, followed by 100,000 further iterations, thinned by 10 so as to obtain 10,000 iterations for use in the probabilistic economic model. Uninformative prior parameters and two different sets of initial values were used; convergence was tested by visual inspection of the Brooks Gelman-Rubin diagram. In addition, convergence of the models was assessed by checking the autocorrelation and the Kernel density plots within WinBUGS. The WinBUGS code used to analyse the relapse data and estimate the underlying Weibull distribution parameters is provided in Table 264. The results of the analysis are shown in Table 265. It can be seen that gamma has a value of less than 1, suggesting that the risk of relapse is reduced over time.

Table 264. WinBUGS code used for synthesis of relapse data in people who are in remission following a single (first) depressive episode, in order to estimate the parameters of the underlying Weibull distribution

Fixed effects model

```
model   {
  for( i in 1 :narms) {
     r[i] ~ dbin(p[i],n[i]) # Binomial likelihood
     p[i] <-1-exp(-lambda*(pow(t[i],gamma))) # Weibull distribution
  }
lambdalog ~ dnorm(0.0,0.1) # vague priors for lambda parameter
log(lambda)<-lambdalog
gammalog ~ dnorm(0.0,0.1) # vague priors for gamma parameter
log(gamma) <-  gammalog
dummy<-s[1]
}
```

Table 265: Results of the data synthesis undertaken in WinBUGS to determine the parameters of the underlying Weibull distribution of the risk of relapse over time, in people who are in remission following a single (first) episode

Parameter	Mean	SD	Median	95% credible intervals
Gamma	0.612	0.057	0.611	0.503 to 0.723
Lambda	0.095	0.010	0.094	0.077 to 0.115

A comparison of the mean modelled cumulative risk of relapse over time (that was utilised in the economic analysis) and the observed cumulative risk of relapse that was extracted from the graphs included in the studies by Eaton et al. (2008) and Mattisson et al. (2007) is provided in Table 266, which suggests that the modelled values are a good approximation of the values observed in the longitudinal studies, taking into account their relative weight in the analysis (the study sample in Mattison et al. (2007) was considerably larger than the study sample in Eaton et al. (2008). The estimated Weibull distribution parameters were used to inform the economic model; more specifically, the time-dependent relapse risk informed the relapse risk in each of the tunnel remission states of the economic model.

Update 2017

Table 266: Cumulative relapse risk over time following remission from a single (first) depressive episode in primary care: modelled and observed risks

Time (years)	Mean modelled risk	Observed risk Eaton et al. (2008)		Observed risk Mattisson et al. (2007)	
		Men [N=22]	Women [N=70]	Men [N=116]	Women [N=228]
1	0.09	0.09	0.06	0.08	0.09
2	0.13	0.14	0.20	0.11	0.13
3	0.17	0.23	0.24	0.14	0.17
4	0.20	0.23	0.27	0.18	0.19
5	0.22	0.23	0.31	0.18	0.22
6	0.25	0.23	0.31	0.20	0.23
7	0.27	0.23	0.37	0.22	0.25
8	0.29	0.23	0.43	0.24	0.27
9	0.30	0.32	0.47	0.26	0.28
10	0.32	0.32	0.50	0.28	0.29

13.2.6.23 Effect of the number of previous depressive episodes on the baseline risk of relapse

There is ample evidence to suggest that the number of previous episodes is a predictor of relapse (Bockting et al., 2006; Hardeveld et al., 2010; Keller & Shapiro, 1981; Kessing & Andersen, 1999; Mueller et al., 1999; Solomon et al., 2000).

Kessing & Andersen (1999) reported the results of a case register study that included all hospital admissions with primary affective disorder in Denmark during 1971–1993. A total of 7,925 unipolar patients were included in the study. The authors reported that the risk of relapse increased with every new episode; the mean hazard ratio of relapse with every additional episode was 1.15 (95% CI 1.11-1.18).

Mueller and colleagues (1999) analysed prospective follow-up data of up to 15 years on the course of major depression for 380 people receiving outpatient specialist care in the US, who recovered from an index episode of major depression. The authors reported a similar mean adjusted odds ratio of relapse for every additional episode of 1.18 (95% CI 1.06–1.31).

The economic model utilised the hazard ratio reported in Kessing & Andersen (1999) in order to estimate the increase in the risk of relapse within each year in remission for every additional depressive episode. Applying this ratio onto the estimated relapse risk for people with one single (no previous) episode allowed estimation of the baseline relapse risk for people with one previous episode and people with three previous episodes (that is, the two populations of interest in the economic analysis). It also allowed estimation of the relapse risk in future remission states (reflecting further previous episodes of relapse) in the model.

The populations in the naturalistic studies that were considered in order to estimate the baseline relapse risk received a range of interventions that were assumed to correspond to clinical management (pill placebo arms) in the economic model. Therefore, the estimated baseline risk of relapse was applied onto the clinical management arms of the economic models, according to the study population (i.e. people having experienced 1 or 3 previous episodes before their 'index' remitted episode).

13.2.7 Probability of remission after relapse

The economic model took into account the chronicity characterising a proportion of depressive episodes. The annual probability of recovery following a relapse of a depressive episode was estimated based on a synthesis of relevant chronicity data included in the review of the naturalistic studies. The GC noted the limited availability of relevant data in

Update 2017

primary care (Table 263). Eaton et al. (2008) reported a probability of persistence of 0.15 over 10 years that suggests a higher chronicity than that observed in secondary care studies; this figure referred to people not remaining free from a depressive episode for at least 1 year, which the GC considered as an unusual criterion for determining chronicity compared with definitions of chronicity in the other studies included in the review. Therefore, this study was not further considered for the estimation of chronicity in the economic model. Riihimäki et al. (2011) reported that the probability of people with depression not reaching full remission in 5 years was 0.30, which is a high figure compared with data on people in primary care reported by Skodol et al (2011) and Stegenga et al (2012). Bukh et al (2016) reported also high chronicity rates compared with other studies in secondary care (Year 1: 0.71; Year 2: 0.42) and was not further considered. In the rest studies included in the review of longitudinal studies, chronicity risks ranged from 0.17-0.33 in the first year (Gonzales et al., 1985; Keller & Shapiro, 1981; Keller et al., 1992; Stegenga et al., 2012); 0.19-0.21 over 2 years (Keller et al., 1984 & 1992), 0.11-0.15 over 3 years (Skodol et al., 2011; Stegenga et al., 2012), 0.12 over 5 years (Holma et al., 2008; Keller et al., 1992), and 0.07 over 10 years (Mueller et al., 1996), which the GC considered a reasonable reflection of the course of depression in clinical practice.

These data suggest that the probability of recovery may also follow a Weibull distribution, with the rate of recovery being higher over the first years of an episode and decreasing with time. As with relapse data, recovery data were synthesised in WinBUGS 1.4.3 using a random effects model (as in this case a larger number of studies on a range of populations from different settings was used), in order to estimate the parameters of the underlying Weibull distribution (lambda and gamma). The model was run with an initial burn-in period of 20,000 iterations, followed by 100,000 further iterations, thinned by 10 so as to obtain 10,000 iterations for use in the probabilistic economic model. Uninformative prior parameters and two different sets of initial values were used; convergence was tested by visual inspection of the Brooks Gelman-Rubin diagram. In addition, convergence of the models was assessed by checking the autocorrelation and the Kernel density plots within WinBUGS. The WinBUGS code used to analyse the recovery data and estimate the underlying Weibull distribution parameters is provided in Table 267. The results of the analysis are shown in Table 268. It can be seen that gamma has a value that is lower than 1, suggesting that the probability of recovery is reduced over time.

Table 267. WinBUGS code used for synthesis of recovery data in people with depression, in order to estimate the parameters of the underlying Weibull distribution

Random effects model

```
model   {
  for( i in 1 :narms) {
     r[i] ~ dbin(p[i],n[i]) # Binomial likelihood
     p[i] <-1-exp(-lambda[s[i]]*(pow(t[i],gamma))) # Weibull distribution
  }
for (j in 1:nstudy){
                    log(lambda[j]) <- lambdalog[j]
                    lambdalog[j]~dnorm(mean.lambdalog,prec.lambdalog) # vague priors for lambda
parameter in each study
}
mean.lambdalog ~ dnorm(0.0,0.1)  # vague priors for mean lambda parameter
prec.lambdalog<-pow(sd.lambdalog,-2)
sd.lambdalog~dunif(0,2)    # precision of mean lambda parameter
log(mean.lambda) <- mean.lambdalog
log(gamma) <-  gammalog # vague priors for gamma parameter
gammalog ~ dnorm(0.0,0.1)
}
```

Update 2017

Table 268: Results of data synthesis undertaken in WinBUGS to determine the parameters of the underlying Weibull distribution of probability of recovery over time, in people in a depressive episode

Parameter	Mean	SD	Median	95% Credible intervals
Gamma	0.440	0.026	0.440	0.389 to 0.491
Mean.lambda	1.171	0.085	1.168	1.016 to 1.344

A comparison of the mean modelled probability of remaining in a depressive episode over time (that was utilised in the economic analysis) and the observed proportions of people remaining in a depressive episode reported in the studies included in the analysis is provided in Table 269, which suggests that the modelled values are a good approximation of the values observed in the longitudinal studies. The estimated Weibull distribution parameters were used to inform the economic model; more specifically, the time-dependent probability of recovery informed each of the tunnel relapse states of the economic model.

Table 269: Probability of remaining in a depressive episode (chronicity) over time: modelled and observed probabilities

Time (years)	Mean modelled probability	Probabilities reported in the literature
0.5	0.39	Stegenga et al., 2012: 0.41; Keller et al., 1992: 0.50
1	0.31	Gonzales et al., 1985: 0.31; Keller & Shapiro, 1981: 0.29; Stegenga et al., 2012: 0.17; Keller et al., 1992: 0.33
2	0.20	Keller et al., 1984: 0.21; Keller et al., 1992: 0.19
3	0.15	Skodol et al., 2011: 0.15; Stegenga et al., 2012 (3.25 years): 0.11
4	0.12	
5	0.09	Holma et al., 2008: 0.12; Keller et al., 1992: 0.12
6	0.08	
7	0.06	
8	0.05	
9	0.05	
10	0.04	Keller et al., 1992 (Mueller et al., 1996): 0.07

Update 2017

13.2.83 Probability of development of side effects from antidepressant treatment

Treatment with antidepressants is associated with the development of various side effects. These can be serious, including death, attempted suicide or self-harm, falls, fractures, stroke or transient ischaemic attack, epilepsy/seizures, myocardial infarction, hyponatraemia and upper gastrointestinal bleeding (Coupland et al., 2011; Jakobsen et al., 2017) or less serious but more common, such as headaches, nausea and other gastrointestinal symptoms, dizziness, agitation, sedation, sexual dysfunction, tremor, sweating, fatigue, and arrhythmia (Anderson et al., 2012; Jakobsen et al., 2017).

Serious side effects from antidepressants are costly to treat and are likely to reduce the quality of life more significantly, in people who experience them. However, they do not occur frequently. Coupland and colleagues (2011) investigated the association between antidepressant treatment and the risk of several potential adverse outcomes in older people with depression, in a retrospective cohort study that utilised data from 60,746 people aged 65 and over diagnosed as having a new episode of depression, obtained across 570 general practices in the UK between 1996 and 2008. The authors reported that SSRIs were associated with the highest adjusted hazard ratios for falls (1.66, 95%; CIs 1.58 to 1.73) and hyponatraemia (1.52; 95% CIs 1.33 to 1.75) compared with when antidepressants were not being used, while a group of 'other antidepressants' defined according to the British National

Formulary, which included mirtazapine and venlafaxine among others, was associated with the highest adjusted hazard ratios for all-cause mortality (1.66; 95% CIs 1.56 to 1.77), attempted suicide or self-harm (5.16; 95% CIs 3.90 to 6.83), stroke/transient ischaemic attack (1.37; 95% CIs 1.22 to 1.55), fracture (1.64; 95% CIs 1.46 to 1.84), and epilepsy/seizures (2.24; 95% CIs 1.60 to 3.15), compared with when antidepressants were not being used. However, for most of these side effects, with the exception of all-cause mortality, the difference in absolute risks between people who received antidepressants and those who did not were small (lower than 1%) with few exceptions: considering the drugs and classes that were included in the guideline economic analysis, for SSRIs, the absolute increase in risk of falls compared with people who did not take antidepressants was 2.21%; for mirtazapine, the absolute increase in risk of attempted suicide or self-harm compared with people who did not take antidepressants was 1.31%. It is noted that these data were derived from older adults with depression, who are likely to have a higher baseline risk for these events compared with younger populations. Therefore, the absolute increase in risk for any of these events in the study population, between those taking antidepressants and those not taking antidepressants, is expected to be lower than that observed between respective groups in older populations.

Jakobsen and colleagues (2017) conducted a systematic review and meta-analysis to assess the effects (including adverse events) of SSRIs versus placebo, 'active' placebo, or no intervention in adult participants with major depressive disorder. The authors reported that SSRIs significantly increased the risks of serious adverse events (odds ratio 1.37; 95% CI 1.08 to 1.75) corresponding to 31/1000 SSRI participants experiencing a serious adverse event compared with 22/1000 control participants (this is a 0.9% difference).

Anderson and colleagues (2012) estimated the prevalence of common side effects such as headaches, nausea or vomiting, agitation sedation and sexual dysfunction associated with treatment with antidepressants, by undertaking a retrospective analysis of data derived from a large US managed care claims form on 40,017 people aged 13 years and above, of whom 36,400 were adults aged 19 years and above, who were newly diagnosed with depression and were initiated on antidepressant monotherapy between 1998 and 2008. Antidepressant groups included, among others, SSRIs, SNRIs, TCAs and tetracyclic antidepressants (which, in 99% of cases, were represented by mirtazapine). The mean time of exposure to antidepressants was 198 days (range 1-2993 days). The authors reported that the most common side effects of those assessed were headaches (ranging from 5.5 to 6.8/1000 person-months of therapy in adults taking one of the above classes of antidepressants) followed by nausea (ranging between 3.6 and 5.5/1000 person-months of therapy in adults taking one of the above classes of antidepressants). The rate of experiencing at least one of the 5 common side effects considered in the study was 9.7/1000 person-months of therapy in adults taking SSRIs, 12.5/1000 person-months of therapy in adults taking SNRIs, 12.6/1000 person-months of therapy in adults taking TCAs and 13.6/1000 person-months of therapy in adults taking mirtazapine. These translate into 11.7, 15.0, 15.2 and 16.3/100 person-years of therapy.

The economic model considered the impact of common side effects on treatment costs and people's HRQoL. A proportion of people receiving SSRIs, TCAs, SNRIs and mirtazapine alone or in combination were assumed to be experiencing common side effects at any time over the duration of maintenance pharmacological treatment. These proportions equalled 0.117 for SSRIs, 0.150 for SNRIs, 0.152 for TCAs and 0.163 for mirtazapine, based on the data reported by Anderson and colleagues (2012). No side effects were considered for people receiving non-pharmacological interventions; however, people receiving non-pharmacological interventions are also expected to experience a range of events such as headaches, nausea or vomiting, etc. The study by Anderson and colleagues (2012) was uncontrolled and did not examine the rate of side effects that were attributable to drugs. Therefore, the economic analysis may have overestimated the impact of common side effects from antidepressants relative to other treatments and thus underestimated their relative cost effectiveness.

Update 2017

The economic model did not incorporate the impact of less common but more severe side effects on costs and people's HRQoL, as this would require most complex modelling and detailed data on the course and management of these side effects. However, omission of these severe side effects is not expected to have considerably affected the results of the economic analysis, due to their low incidence in the study population. Nevertheless, omission of less common but severe side effects from the economic analysis may have potentially overestimated the cost effectiveness of pharmacological and combined treatments.

13.2.98 Mortality

Depression is associated with an increased risk of mortality relative to the general population. A comprehensive systematic review of 293 studies that assessed the increased risk of people with depression relative to non-depressed individuals, which included 1,813,733 participants (135,007 depressed and 1,678,726 non-depressed) reported a risk ratio of mortality in depressed relative to non-depressed participants of 1.64 (95% CI 1.56 to 1.76). After adjustment for publication bias, the overall risk ratio was reduced to 1.52 (95% CI 1.45 to 1.59) (Cuijpers et al., 2014). The adjusted figure was applied onto general mortality statistics for the UK population (ONS, 2015), to estimate the absolute annual mortality risk in people experiencing a depressive episode relative to people not experiencing a depressive episode within each cycle of the model. People with a depressive episode were assumed to be at increased mortality risk due to depression only in the years they experienced a depressive episode (i.e. while they were in the relapse health state). The same mortality risk was assumed for both men and women experiencing a relapse, as no gender-specific data were reported in the study. People not experiencing a depressive episode in each model cycle were assumed to carry the mortality risk of the general UK population.

It is acknowledged that the mortality risk ratio refers to depressed versus non-depressed individuals and not versus the general population. The UK general population already includes a proportion of people with major depression: according to the latest adult psychiatric morbidity survey for England, 3.3% of adults suffered from depression in 2014 (McManus et al., 2016); therefore the economic analysis has slightly overestimated the annual mortality risk for people experiencing a depressive episode as well as for those not experiencing a depressive episode. This is a limitation of the analysis owing to lack of more appropriate data, which, nevertheless, is expected to have had a negligible effect on the cost effectiveness results.

Update 2017

13.2.103 Utility data and estimation of quality adjusted life years (QALYs)

In order to express outcomes in the form of QALYs, the health states of the economic model need to be linked to appropriate utility scores. Utility scores represent the health-related quality of life (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 review of utility data on depression-related heath states identified 5 studies that reported utility data corresponding to depression-related health states, which were derived from EQ-5D measurements on adults with depression valued by the general UK population (Kaltenthaler et al., 2006; Koeser et al., 2015; Mann et al., 2009, Sapin et al., 2004; Sobocki et al., 2006 & 2007). Three of the studies analysed EQ-5D data obtained from adults with depression or common mental health problems participating in RCTs conducted in the UK (Kaltenthaler et al., 2006; Koeser et al., 2015; Mann et al., 2009). The other two studies analysed naturalistic primary care EQ-5D data from adults with depression in France (Sapin et al., 2004) and in Sweden (Sobocki et al., 2006 & 2007). All studies reported utility values associated with severity of depression (e.g. mild, moderate or severe) and/or states of depression relating to treatment response (e.g. response, remission, no response) and were thus relevant to the health states considered in economic modelling conducted for this

guideline. All studies defined health states using validated measures of depressive symptoms, such as the BDI, the HAMD-17, the PHQ-9, the MADRS and the CGI.

An overview of the study characteristics, the methods used to define health states, and the health-state utility values reported by each of the studies is provided in Table 270.

All reported utility data comply with the NICE criteria on selection of utility data for use in NICE economic evaluations (NICE, guide to methods for TA 2013). The data from Kaltenthaler and colleagues (2006) were derived following mapping of CORE-OM data onto BDI data; however, the BDI cut-off scores used to determine the health states by depressive symptom severity were not reported, and therefore it is not clear the exact level of symptom severity the resulting utility scores correspond to. All other studies provided details on the scale cut-off scores used to determine the depression-related health states by severity or by response to treatment. Mann and colleagues (2009) used the original PHQ-9 cut-off scores to determine severity levels of depression. However, it is noted that a PHQ-9 score of 5-9, which corresponded to the state of mild depression according to the PHQ-9 manual, is also below the cut-off point for clinically detected depression (Gilbody et al., 2007a & 2007b).

The economic model of interventions aiming at relapse prevention used data from Sobocki and colleagues (2006 & 2007). This was decided because the study provided data that could be linked to all states included in the model, i.e. relapse to less severe depression (the value of 0.60 for mild depression was used), relapse to more severe depression (a weighted average of the utility of moderate and severe depression of 0.42 was used) and remission (0.81) and was based on a larger study sample compared with the rest studies providing utility data. Remission was defined in the study as an improved or very much improved score on the CGI-Improvement scale, combined with a clinical judgement by the treating doctor of being in full remission. It is acknowledged that this definition of remission may actually indicate response to treatment not reaching full remission. Nevertheless, although all cohorts enter the model in full remission, a proportion of people in the cohorts remitting from future episodes might not experience full remission and might have some residual symptoms, and therefore the utility value of remission based on the improved or very much improved CGI-I score is likely to express the utility of people in future remission states. It is noted that the value of 0.81 corresponding to the state of ‘remission’ in Sobocki and colleagues (2006 & 2007) is very close to the utility value of remission (0.80) reported in Koeser and colleagues (2015) and between the values of 0.72 and 0.85 corresponding to the states of ‘response not reaching remission’ and ‘response reaching remission’, respectively, that were reported by Sapin and colleagues (2004) (who defined response and remission based on MADRS scores), which indicates that the value utilised in the model may reflect a utility between partial and full remission that is closer to the utility of the latter.

For people relapsing to less severe depression and more severe depression the higher values of 0.65 and 0.56, respectively, reported in Mann and colleagues (2009) were tested as a more conservative scenario in sensitivity analysis.

Update 2017

Table 270: Summary of available EQ-5D derived health-state utility data for depression (UK tariff)

Study	Definition of health states	Health state / severity	N	Mean (SD or 95% CI)
Kaltenthaler et al., 2006	Analysis of EQ-5D and CORE-OM data obtained 62 people with common mental health problems participating in a multi-centre RCT of supervised self-help CBT in the UK (Richards et al., 2003). CORE-OM data were first mapped onto the BDI, which was used to categorise people into 3 groups of mild to moderate, moderate to severe and severe depression. BDI cut-off scores used for categorisation were not reported. EQ-5D utility value for no depression obtained from age- and gender-matched normal population in the UK (Kind et al., 1998).	No depression Mild to moderate Moderate to severe Severe	NR NR NR NR	0.88 (0.22) 0.78 (0.20) 0.58 (0.31) 0.38 (0.32)
Koeser et al., 2015	Analysis of EQ-5D and HAMD17 data obtained from people with recurrent depression in full or partial remission participating in a RCT of MBCT in the UK (N=123) (Kuyken et al., 2008). Definition of health states by HAMD scores: remission ≤ 7; response 8-14; no response ≤ 15	Remission Response No response	NR NR NR	0.80 (0.02) 0.62 (0.04) 0.48 (0.05)
Mann et al., 2009	Analysis of EQ-5D and PHQ-9 data collected from 114 people with depression participating in a cluster RCT of collaborative care across 19 UK primary care practices based in urban and rural communities (Richards et al., 2008). Definition of health states by PHQ-9 score: mild 5-9; moderate 10-14; moderately severe 15-19; severe 20-27	Mild Moderate Moderate to severe Severe	10 24 39 35	0.65 (0.23) 0.66 (0.21) 0.56 (0.27) 0.34 (0.29)
Sapin et al., 2004	Analysis of EQ-5D and MADRS data collected from 250 people with major depression recruited from 95 French primary care practices for inclusion in an 8-week follow-up cohort. Definition of health states by MADRS score: remission MADRS ≤ 12; response at least 50% reduction in the MADRS baseline score over 8 weeks. Baseline mean MADRS score 32.7 (SD 7.7)	Response – remission Response – no remission No response Baseline	144 34 46 250	0.85 (0.13) 0.72 (0.20) 0.58 (0.28) 0.33 (0.25)
Sobocki et al., 2006 & 2007	Analysis of EQ-5D and CGI-S and CGI-I data collected from 447 adults with depression enrolled in a naturalistic longitudinal observational 6-month study conducted in 56 primary care practices in 5 regions of Sweden. People who started a new or changed antidepressant treatment were eligible for inclusion. Definition of health states by CGI-S score: mild 2-3; moderate 4; severe 5-7; remission 'much or very much improved' score (1-2) combined with clinical judgement	Mild Moderate Severe Remission No remission	110 268 69 207 191	0.60 (0.54 to 0.65) 0.46 (0.30 to 0.48) 0.27 (0.21 to 0.34) 0.81 (0.77 to 0.83) 0.57 (0.52 to 0.60)
Notes: CI: confidence intervals; N: number of participants who provided ratings on the EQ-5D; NR: not reported; SD: standard deviation				

Update 2017

According to the GC expert opinion, an average depressive episode lasts 6 months. This estimate is supported by data from a prospective study on 250 adults with a newly originated (first or recurrent) major depressive episode, drawn from a prospective epidemiological Dutch survey on 7,046 people in the general population (Spijker et al., 2002). According to this study, the mean duration of a recurrent episode was 6.1 months (95% CI 4.7-7.5). The economic model assumed that people experiencing a depressive episode that resolved in the next year (i.e. people who spent only a year in the depressive episode and then moved to the remission state in the next cycle), experienced a reduction in their HRQoL for 6 months out of the 12 months of the cycle they remained in the 'relapse' state. Thus, people relapsing to depressive episodes that lasted only for one year were assumed to have the utility of remission for 6 months and the utility of depression (less or more severe) for another 6 months. However, people whose depressive episode lasted for at least 2 cycles (years) were attached the utility of depression over the number of years they remained in relapse except their final year in the relapse state, in which they were assumed to have the utility of depression for 6 months and the utility of remission for another 6 months.

Side effects from medication are expected to result in a reduction in utility scores of adults with depression. Sullivan and colleagues (2004) applied regression analysis on EQ-5D data (UK tariffs) obtained from participants in the 2000 national US Medical Expenditure Panel Survey to derive age-adjusted utility values for health states associated with depression and with side effects of antidepressants. Health states were defined based on descriptions in the International Classification of Diseases (9th Edition) [ICD-9] and the Clinical Classification Categories (CCC) [clinically homogenous groupings of ICD-9 codes derived by the Agency for Healthcare Research and Quality]. Table 271 shows the health states determined by Sullivan and colleagues (2004) and the corresponding utility values obtained from regression analysis of EQ-5D data. The mean utility decrements due to side effects from antidepressants ranged from -0.044 (diarrhoea) to -0.129 (excitation, insomnia and anxiety), with a mean decrement of -0.087. This mean utility decrement was applied to the proportion of people who experienced side effects from maintenance antidepressant treatment alone or in combination, over the whole duration of antidepressant treatment, i.e. over 2 years.

Update 2017

Table 271: Summary of EQ-5D derived health-state utility data for side effects from antidepressants (UK tariff)

Study	Definition of health states	Health state	Mean (95% CI)
Sullivan et al., 2004	Censored least absolute deviations (CLAD) regression analysis of EQ-5D data from the 2000 national US Medical Expenditure Panel Survey (MEPS) [http://meps.ahrq.gov/mepsweb/] Definitions of health states Gastrointestinal symptoms (GI): average Diarrhoea: clinical classification categories (CCC) - Agency for Healthcare Research and Quality): 144 regional enteritis Dyspepsia: CCC 138 oesophageal disorders Nausea & constipation: assumed average of GI Sexual: ICD-9 302 sexual disorders Excitation: average Insomnia: assumed equal to anxiety Anxiety: CCC 072 anxiety, somatoform, dissociative disorders Headache: CCC 084 headache Drowsiness & other: assumed average of all side effects Untreated depression ICD-9 311 depressive disorder; CLAD 25% Treated depression: ICD-9 311 depressive disorder; CLAD 75%; baseline utility estimate (not a decrement)	GI symptoms Diarrhoea Dyspepsia Nausea Constipation Sexual Excitation Insomnia Anxiety Headache Drowsiness Other Untreated depression Treated depression	-0.065 (-0.082 to -0.049) -0.044 (-0.056 to -0.034) -0.086 (-0.109 to -0.065) -0.065 (-0.082 to -0.049) -0.065 (-0.082 to -0.049) -0.049 (-0.062 to -0.037) -0.129 (-0.162 to -0.098) -0.129 (-0.162 to -0.098) -0.129 (-0.162 to -0.098) -0.115 (-0.144 to -0.087) -0.085 (-0.107 to -0.065) -0.085 (-0.107 to -0.065) -0.268 (-0.341 to -0.205) 0.848 (0.514 to 0.971)

Update 2017

13.2.11 Resource use – intervention costs

Intervention costs were estimated by combining resource use associated with each intervention with appropriate unit costs (drug acquisition costs and healthcare professional unit costs).

13.2.11.1 Maintenance pharmacological treatment

Pharmacological intervention costs consisted of drug acquisition and GP visit costs. In addition to the 3 class-representative drugs (citalopram for SSRIs, venlafaxine for SNRIs, amitriptyline for TCAs) and mirtazapine, the model also considered clinical management (reflected in the placebo arms of the relapse prevention RCTs), which comprised GP visits only. The cost of fluoxetine maintenance treatment was also estimated, as fluoxetine was considered as part of combined pharmacological and psychological maintenance treatment, as well as a separate treatment option, in people who remitted following combination or psychological therapy.

The average daily dosage for each drug was determined according to optimal clinical practice (BNF 2016), following confirmation by the GC in order to reflect routine clinical practice in the NHS, and was consistent with dosages reported in the RCTs that were included in the systematic review of interventions for relapse prevention in adults with depression.

Maintenance pharmacological treatment lasted 2 years, based on available relevant evidence and previous NICE guidance. The model assumed gradual discontinuation (tapering) of the drug at the end of maintenance treatment, which was modelled as a linear reduction of the drug acquisition cost (from optimal dose to zero) in the last month of maintenance treatment, according to routine clinical practice, as advised by the GC.

Provision of maintenance pharmacological treatment involved 6 GP contacts in the 1st year of treatment and another 3 in the 2nd year; one extra GP visit was assumed during the tapering period. Clinical management (placebo) comprised 3 GP contacts in the 1st year and 1 contact in the 2nd year of treatment. For people in remission following pharmacological treatment who subsequently received clinical management as maintenance treatment option, a tapering period in the first month of the intervention was assumed, which included a month of antidepressant administration in a linearly reduced dose (starting from optimal dose until no drug was received) plus one extra GP visit.

These resource use estimates were based on the GC expert advice; they represent UK optimal routine clinical practice but may be lower than some of the descriptions of medical resource use in pharmacological trial protocols, where resource use is more intensive than clinical practice.

The drug acquisition costs and the GP unit cost were taken from national sources (National drug tariff January 2017, Curtis & Burns, 2016). The lowest reported price for each drug was used, including prices of generic forms, where available. The reported GP unit cost included remuneration, direct care staff costs and other practice expenses, practice capital costs and qualification costs. The latter represented the investment costs of pre-registration and postgraduate medical education, annuitised over the expected working life of a GP; ongoing training costs were not considered due to lack of available information. The unit cost per patient contact was estimated taking into account the GPs' working time as well as the ratio of direct (surgeries, clinics, telephone consultations & home visits) to indirect (referral letters, arranging admissions) patient care, and time spent on general administration.

Intervention costs of maintenance pharmacological treatment and of clinical management (reflected in treatment with placebo) are shown in Table 272.

Update 2017

Table 272: Intervention costs of maintenance pharmacological treatments considered in the guideline economic analysis on relapse prevention (2016 prices)

Drug	Mean daily dosage	Drug acquisition cost[1]	2-year drug cost (includes one month tapering)	2-year total intervention cost (drug and GP[2])
Citalopram	50% 20mg 50% 45mg	20mg, 28 tab, £0.83 40mg, 28 tab, £1.01	£23.24	£383.24
Venlafaxine	150mg in 2 doses	75mg, 56 tab, £2.19	£55.92	£415.92
Amitriptyline	75mg	25mg, 28 tab, £0.79	£60.52	£420.52
Mirtazapine	50% 30mg 50% 45mg	30mg, 28 tab, £1.27 45mg, 28 tab, £1.55	£36.01	£396.01
Fluoxetine[3]	20mg	20mg, 30 cap, £0.87	£20.74	£380.74
Placebo (clinical management)	Linear reduction over 1 month	As above, depending on tapered acute drug treatment (if applicable)	£0-£11.20[4]	£144.00[5]-£181.27

[1] (national drug tariff, January 2017)

[2] GP cost includes 6 GP visits in the 1st year and 3 GP visits in the 2nd year, plus a visit during tapering (GC expert opinion); GP unit cost £36 per patient contact lasting 9.22 minutes (Curtis & Burns, 2016)

[3] Fluoxetine was considered as part of combination maintenance treatment, as well as a separate treatment option, in people who remitted following combined treatment.

[4] depends on whether tapering is required (i.e. whether acute treatment was pharmacological and which drug was used); range of drug cost reflects range of drug acquisition cost during tapering

[5] lower estimate does not include tapering visit

Update 2017

13.2.11.23 Maintenance psychological interventions

Maintenance psychological therapies comprised a number of individual or group sessions delivered by a range of healthcare professionals. Resource use estimates of each maintenance psychological therapy in terms of number and duration of sessions, mode of delivery and number of therapists and participants in the case of group interventions were determined by resource use data described in respective RCTs that were included in the guideline systematic review, confirmed by the GC to represent clinical practice in the UK; where trial resource use was very different to routine UK practice, a sensitivity analysis was undertaken, testing the impact of using routine UK resource use estimates on the results of the analysis. Unit costs were taken from national sources and were assumed to correspond, on average, to an Agenda for Change (AfC) band 7 clinical psychologist, as expressed in MBCT therapist costs (Curtis & Burns, 2016). The reported therapist unit costs included wages/salary, salary oncosts, capital and other overheads, but no qualification costs.

Qualification costs for clinical psychologists were obtained from a separate source (National College for Teaching and Leadership, NHS Health Education England, 2016). According to this, the average cost of training a clinical psychology trainee reaches £159,420 over 3 years, comprising £49,074 of tuition fees, £107,073 of salary (including on-costs) and £3,273 of placement fees (2016 prices). Using a working life of a clinical psychologist of 25 years (according to GC expert advice), the annuitized qualification cost of clinical psychologist was estimated at £9,673.

The GC also advised that delivery of MBCT by clinical psychologists requires extra training that is not included in qualification costs. This training cost has been estimated to reach £1,500 per trainee, based on expert advice. Using a higher estimate of £3,000 per trainee, assuming that this is a one-off training cost and that the therapist has a working life of 25

years, the annuitised training cost specific to MBCT is £176. Assuming a conservative annual volume of MBCT clients of 30 per therapist, then the training cost associated with MBCT is £6 per client. This cost is trivial and is likely to be even lower due to deliberately high figure used for the overall training cost, and the conservative figures used for the working life of a MBCT therapist and the annual volume of MBCT clients per therapist. Therefore, this cost was not considered further when calculating the unit cost of a therapist delivering MBCT.

Ongoing training costs of clinical psychologists were also not considered, because no relevant data are available. It is noted that this approach is consistent with the lack of consideration of ongoing training costs in the estimation of the reported GP unit cost, also due to lack of relevant data.

The GC also advised that supervision costs be considered in the estimation of the clinical psychologist unit cost, as supervision is essential for the delivery of psychological therapies and may incur considerable costs. According to the British Association for Behavioural and Cognitive Therapies, therapists should receive regular supervision in groups of no more than 6 participants, with a mean duration of 1.5 hour per month for a full time practitioner (British Association for Behavioural and Cognitive Therapies, 2016). According to expert advice, MBCT therapists should receive approximately an hour of supervision per month, by a NHS Band 7 or 8 supervisor, sometimes offered in groups of 2-4 therapists. Based on this information, supplemented with GC expert advice, the same annual supervision cost was estimated for both CBT/CT and MBCT therapists, comprising 1 hour of supervision per month, delivered by a Band 8a (AfC) clinical psychologist in groups of 4 therapists. The estimated annual supervision cost per supervised therapist and details considered for its estimation are provided in Table 273. This supervision cost includes the cost of the supervisor's time, but not the cost of the supervised therapist's time, as this is indirectly included in the unit cost of a clinical psychologist, as discussed below.

Table 273: Annual cost of supervision for therapists delivering CBT/CT or MBCT (2016 prices)

Cost element	Unit cost (annual)	Source
Wages – salary	£46,095	Curtis & Burns, 2016; unit cost of community-based scientific and professional staff (Agenda for Change band 8a)
Salary on-costs	£11,702	
Overheads – staff	£14,160	
Overheads - non-staff	£22,079	
Capital overheads	£4,583	
Qualifications	£9,673	Based on a mean clinical psychologist training cost estimate of £159,420 (National College for Teaching and Leadership, NHS Health Education England, 2016) and a working life of 25 years
Total cost	£108,292	
Working time	42.4 weeks /year 37.5 hours /week (1,590 hours)	Curtis & Burns, 2016
Total cost per hour	**£68**	
Annual cost of supervision of group of 4 therapists (reflecting supervisor's time spent on supervision)	£1,226	Based on 1.5 hour supervision per month (British Association for Behavioural and Cognitive Therapies, 2016 and expert advice)
Annual supervision cost per supervised therapist	**£306**	Based on delivery of supervision in groups of 4 participants (British Association for Behavioural and Cognitive Therapies, 2016 and expert advice)

Update 2017

In estimating the unit cost of a clinical psychologist per hour of client contact, the ratio of direct (face-to-face) to indirect time (reflecting time for preparation of therapeutic sessions and other administrative tasks) of a clinical psychologist was taken into account. According to GC expert opinion, delivery of individual therapies lasting 1 hour requires 15 minutes of preparation, whereas delivery of group therapies lasting 2 hours requires 30 minutes of preparation time. This results in a ratio of direct to preparation time of 1: 0.25, which is independent of the mode of delivery of psychological interventions; this ratio does not take other administrative tasks (that increase the therapist's indirect time) into account. In MBCT trials conducted in the UK, the ratio of direct to indirect time of MBCT therapists has been reported to equal 1: 0.67 (Kuyken et al., 2008 & 2015-HTA); this estimate, however, was based on the time of 3 therapists. Curtis and Burns (2016) report a 1: 1 direct to indirect time ratio for CBT therapists delivering services for children and young people, based on information from a trial of SSRIs with or without CBT in adolescents with major depression. Curtis (2014) reports a 1: 1.25 direct to indirect time ratio for clinical psychologists based on the National Child and Adolescent Mental Health Service mapping data and returns from over 500 principal clinical psychologists, but it is acknowledged that this level of seniority may involve more supervision and managerial time, so the ratio may be an overestimate of the direct to indirect time of a AfC Band 7 clinical psychologist. After reviewing this information on the ratio of direct to indirect time of clinical psychologists, the GC advised that the direct to indirect ratio of a therapist of Band 7 delivering CBT/CT or MBCT is 1: 0.67 and this ratio was utilised in the economic model.

An overview of the cost elements that were taken into account in the estimation of the unit cost of a clinical psychologist delivering psychological therapies in the economic model is shown in Table 274.

Table 274: Unit cost of clinical psychologist (2016 prices)

Cost element	Unit cost (annual)	Source
Wages – salary	£38,173	Curtis & Burns, 2016; unit cost of MBCT therapist (Agenda for Change band 7)
Salary on-costs	£9,500	
Overheads – staff	£11,680	
Overheads - non-staff	£18,211	
Capital overheads	£4,583	
Qualifications	£9,673	Based on a mean clinical psychologist training cost estimate of £159,420 (National College for Teaching and Leadership, NHS Health Education England, 2016) and a working life of 25 years
Supervision	£306	See Table 273 for details
SUM of unit costs	**£92,126**	
Working time	42.4 weeks /year 37.5 hours /week (1,590 hours)	Curtis & Burns, 2016
Total cost per hour	**£58**	
Ratio of direct to indirect time[1]	1:0.67	Curtis & Burns, 2016; assumption based on GC expert opinion and a review of respective ratios reported in the literature for clinical psychologists and other therapists delivering psychological interventions
Estimated cost per hour of direct contact	**£97**	

1 ratio of face-to-face time to time for preparation and other administrative tasks

Update 2017

In addition, according to the GC expert advice, people receiving maintenance psychological therapy had 2 contacts with a GP during maintenance treatment.

Details on resource use and total costs of maintenance psychological interventions are provided in Table 275.

Table 275: Intervention costs of maintenance psychological therapies considered in the guideline economic analysis on relapse prevention (2016 prices)

Intervention	Resource use details	Total intervention cost per person[1]
MBCT	8 group sessions + 4 group booster sessions lasting 2 hours each; 1 therapist and 12 participants per group = 24 therapist hours per group and 2 therapist hours per service user	£193 + £72
CBT	10 individual sessions lasting 1 hour each	£966 +£72
CT	10 individual sessions lasting 1 hour each	£966 +£72
Group CT	8 group sessions lasting 2 hours each; 1 therapist and 10 participants per group = 16 therapist hours per group and 1.6 therapist hours per service user	£155 +£72
1 cost of psychological intervention plus 2 GP visits, at a GP unit cost £36 per patient contact lasting 9.22 minutes (Curtis & Burns, 2016); cost of psychological intervention based on resource use combined with unit cost of therapist per hour of direct contact with client, estimated as described in Table 274.		

The GC considered the resource use associated with individual CBT and CT (Table 275) to be substantially higher than the level of intensity of maintenance psychological treatment received in routine UK practice. For this reason a sensitivity analysis was carried out that tested the impact of reducing the number of individual CBT or CT sessions down to 4, on the results of the economic analysis.

Update 2017

13.2.11.3 Combined maintenance pharmacological and psychological intervention

The intervention cost of combined maintenance pharmacological and psychological intervention was estimated as the sum of the intervention costs of the individual pharmacological and psychological treatment components.

In cohorts receiving combination treatment, no extra GP visits were added onto the psychological intervention cost, since people were already receiving GP care as part of their antidepressant treatment.

13.2.12 Cost of relapse and remission states

The cost of relapse and remission states in the economic model was estimated based primarily on data from Byford et al. (2011). This was a naturalistic, longitudinal study that aimed to estimate the health service use and costs associated with non-remission in people with depression using data from a large primary care UK general practice research database between 2001 and 2006. The study analysed 12-month healthcare resource use data on 88,935 adults with depression and in receipt of at least two antidepressant prescriptions (for amitriptyline, citalopram, escitalopram, fluoxetine, paroxetine, sertraline or venlafaxine) in the first 3 months after the index prescription. The study provided data on resource relating to medication (antidepressant use and concomitant medication such as anxiolytics, hypnotics, mood stabilizers and neuroleptics), GP contacts, psychological therapy, psychiatrist and other specialist contacts, inpatient stays and accident and emergency attendances. Data were reported separately for people who remitted within 12 months, and those who did not remit. In addition, the study included graphs showing the change in healthcare costs overtime by timing of remission (separate graph lines were provided for people with very early remission defined as 1-4 months after onset of the depressive episode, early remission

occurring 5-9 months after onset of the episode, late remission occurring 9-12 months after onset of the depression episode, and for people not achieving remission by 12 months). According to the study, among study participants who successfully ceased antidepressant treatment within the first 12 months (most probably remitters), 40% ceased within 4 months of the index prescription and almost 80% ceased within 8 months. This suggests that the costs incurred after remission did not include maintenance pharmacological treatment costs but were instead healthcare costs unrelated to depression.

Healthcare resource use and cost data from this study were modified following GC advice and attached to the model health states: data on people in a depressive episode who remitted within 12 months in the study were attached onto people in the relapse state of the model in their final year before remission, and also to people whose depressive episode lasted only over one model cycle. Resource use and cost data on people who did not remit within 12 months in the naturalistic study were used as the basis for estimating healthcare costs incurred by people who remained in a depressive episode for longer than one year and were applied to all years in a relapse state except the year before remission. Costs incurred after remission was achieved (which were possible to obtain from the graphs using digital software) were used to estimate annual healthcare costs associated with the remission state of the model.

Following GC advice, some of the resource use and drug acquisition cost data reported in the paper were modified, to reflect current clinical practice and the fact that some drugs are now available off patent. Some cost data were sought from other sources. Where detailed resource use data were provided, these were combined with appropriate 2016 unit costs; where only cost figures were available, these have been uplifted to 2016 prices using the hospital and community health services (HCHS) index (Curtis & Burns, 2016), so that all costs in the guideline economic analysis reflect 2016 prices.

The resource use and cost data reported in the paper by Byford and colleagues (2011) for people with depression who remitted and those who did not remit within 12 months from the index prescription, uplifted to 2016 prices using the HCHS index, are presented in Table 276.

Update 2017

Table 276: Reported 12 month resource use and costs reported in Byford and colleagues (2011) (cost figures uplifted to 2016 prices)

Resource use element	Remitters (n=53,654)					Non-remitters (n=35,281)				
	Resource use			Cost		Resource use			Cost	
	Use %	Mean	SD	Mean	SD	Use %	Mean	SD	Mean	SD
Antidepressant use				£82	£54				£190	£84
Number of prescriptions	100	4.8	3.2			100	11.1	5.7		
Cumulative duration (days)		155.2	101.5				358.7	158.4		
Time on treatment (days)		129.8	73.7				283.9	63.8		
Concomitant medication				£33	£168				£80	£335
Anxiolytics – BZD (days)	8.2	32.4	241.7			12.6	69.5	458.5		
Anxiolytics – other (days)	0.7	0.8	15.0			1.1	1.6	23.7		
Hypnotics – BZD (days)	11.4	39.8	258.7			16.9	84.0	552.1		
Hypnotics – Z drugs (days)	9.2	7.5	44.4			12.9	16.4	71.6		
Hypnotics – other (days)	0.5	0.8	22.1			0.6	1.5	30.3		
Mood stabilizers – Li (days)	1.2	6.0	47.9			3.1	12.7	90.2		
Mood stabilizers – antiepileptic (days)	4.7	2.2	31.5			6.2	8.5	72.4		
Neuroleptics – typical (days)	0.2	0.4	11.2			0.5	1.4	25.9		
Neuroleptics – atypical (days)	0.7	3.0	54.8			1.1	8.3	120.0		
Service use										
GP visits	100	12.9	8.9	£436	£300	100	17.3	10.4	£619	£345
GP phone calls	55.2	2.5	4.3			86.7	5.4	6.1		
Psychological therapy contacts	0.2	0.0	0.1	£0	£4	0.2	0.0	0.1	£0	£8
Psychiatrist contacts	2.9	0.0	0.3	£89	£154	5	0.1	0.4	£115	£184
Other specialist contacts	38.6	0.6	1.1			44.9	0.8	1.2		
Hospitalisations [admissions]	5.2	0.1	0.4	£163	£847	5.7	0.1	0.4	£190	£982
Accident and emergency attendances	3.1	0.0	0.3	£6	£37	3.3	0.1	0.3	£6	£37
TOTAL COST				**£809**	**£1,044**				**£1,200**	**£1,252**

Update 2017

Costs for each healthcare cost category associated with the treatment of people with depression who remitted and those who did not remit within 12 months from their index episode were estimated as follows:

Cost of antidepressants and concomitant medication – relapse and remission states

The GC noted that a number of antidepressant drugs have become generic since the time the study was conducted, and this would have resulted in a reduction in the antidepressant costs reported in the study. In order to attach up-to-date drug acquisition costs to the antidepressant use reported in the study for 2001-2006, the following methodology was used: based on national prescription cost data for England in 2006 and 2015 - the most recent year for which relevant data existed - (NHS, The Information Centre 2007; Prescribing & Medicines Team, Health and Social Care Information Centre, 2016), the ratio of the net ingredient cost (NIC) per antidepressant prescription item of 2015 relative to 2006 (which was the cost year used in the study by Byford and colleagues) was calculated; this was 0.50 (NIC per antidepressant prescription item was 9.39 in 2006 and 4.67 for 2015), and suggests that the mean cost per prescription has been reduced by 50%. Subsequently, the mean acquisition cost of antidepressants in 2015 was adjusted to be 50% lower than the cost reported in 2006.

Similarly to the methodology described above, for each category of concomitant medication, the ratio of the NIC per prescription item of 2015 relative to 2006 was calculated, and this was applied as a weighted ratio (according to the concomitant medication usage reported in the study) onto the cost of concomitant medication reported in the study, to adjust the total cost of concomitant medication to 2015 price.

The NICs per prescription items for antidepressants and the broad categories of concomitant medication in years 2006 and 2015 as well as the resulting ratios of 2015:2006 NICs are provided in Table 277.

Table 277: Net ingredient cost (NIC) per prescription item for antidepressants and categories of concomitant medication in 2006 and 2015

Drug category	NIC 2006	NIC 2015	Ratio NIC 2015:2006
Antidepressants	9.39	4.67	0.50
Anxiolytics	3.66	2.36	0.64
Hypnotics	2.75	6.78	2.47
Mood stabilizers – Li carbonate	1.72	1.50	0.87
Mood stabilizers – antiepileptic	21.54	22.79	1.06
Neuroleptics	38.83	13.69	0.35
Source: NHS, The Information Centre 2007; Prescribing & Medicines Team, Health and Social Care Information Centre, 2016			

Byford and colleagues (2011) reported that among study participants who successfully ceased antidepressant treatment within the first 12 months (most probably remitters), 40% ceased within 4 months of the index prescription and almost 80% ceased within 8 months. On the other hand, among participants who did not meet criteria for remission, 60% discontinued antidepressant treatment at some point over the 12-month study period but resumed within 6 months of antidepressant cessation and 40% received continuous antidepressant treatment over the 12-month study period.

Following GC expert opinion and previous NICE guideline recommendations on optimal duration of maintenance antidepressant treatment after remission of a depressive episode, the economic model assumed that antidepressant treatment for each depressive episode lasted in total for at least 2 years; more specifically, it lasted over the duration of the depressive episode (i.e. over the whole period people spent in a relapse state) plus the first

Update 2017

year into remission. Therefore, the adjusted estimated 12-month antidepressant cost for remitters was applied to all remitters in the model over their first year of remission, to reflect continuation of maintenance pharmacological treatment according to NICE guidance.

GP visits and phone contacts – relapse and remission state

To estimate associated costs, relevant resource use for remitters and non-remitters reported in Byford and colleagues (2011) was combined with respective unit costs (Curtis and Burns, 2016).

Moreover, 3 extra GP visits were estimated for those who remitted in their first year of remission, to reflect extra resource use and costs associated with maintenance pharmacological treatment.

Cost of psychological therapy – relapse state

The GC noted that the study by Byford and colleagues (2011) reported a very low usage of psychological therapies. This is attributable to two reasons: first, because people in the study were selected for receiving antidepressant therapy, and second, because psychological therapy was not widely offered at the time the study was conducted (which was prior to the establishment of the IAPT programme in the UK).

According to NHS England, IAPT end of year data suggested that the percentage of people referred to IAPT services and receiving psychological therapies among those presenting to their GP and being eligible for psychological treatment reached 16.8% in 2016 (NHS England, 2016).

Radhakrishnan et al (2013) reported costs of IAPT services in 5 East of England region Primary Care Trusts. Costs were estimated using treatment activity data and gross financial information, along with assumptions about how these financial data could be broken down. Data referred to 8,464 clients who attended at least 2 sessions (of whom 4,844 completed treatment). Using baseline PHQ-9 score bands to assess severity of depression, 2146 patients (25.4%) were classified as having moderate depressive symptoms, 1987 patients (23.5%) had moderate-severe depressive symptoms and 1787 patients (21.1%) presented with severe depressive symptoms. Based on the data reported in the study, the weighted mean cost per course of IAPT treatment per person (including people who completed treatment, those who dropped out, people who declined treatment and also people who were judged not to be suitable for treatment) was estimated to reach £740 (2016 prices). This unit cost was multiplied by the percentage of people receiving psychological therapy to estimate the cost of psychological treatment in the economic cohort, which was added to the annual cost of both people who remained in the relapse state, and those who moved to remission in the next model cycle.

The GC advised that people receiving psychological therapy still have GP contacts and some may also receive combination therapy. Therefore the costs of psychological treatment were added to the total cost associated with the relapse state, without other costs being reduced.

Cost of secondary care – relapse state

The cost of hospitalisation, psychiatrist visits, visits to other specialists and accident and emergency attendances was estimated by multiplying relevant resource use reported in Byford and colleagues (2011) by respective NHS reference unit costs (Department of Health, 2016).

For hospitalisation, the mean cost per elective admission in NHS care was used. The GC expressed the opinion that a proportion of hospitalisations in the cohort should be due to their depressive episode. However, this proportion was not possible to estimate. Therefore the GC decided to use the mean total cost per admission in the NHS as a conservative

Update 2017

estimate of the cost of hospitalisation (since admissions to psychiatric wards are more expensive).

Cost of remission state

According to the graphs presented in the Byford et al. (2011) study, the data of which were possible to extract using digital software (http://www.digitizeit.de), the 3-month costs after people had reached remission were approximately £100, thus the annual costs of remission reached £400 (2006 prices). Since the paper reports that over 40% of participants who successfully ceased antidepressant treatment ceased within 4 months of the index prescription and almost 80% ceased within 8 months, this cost figure appears not to be associated with maintenance treatment of the depressive episode, but is rather a 'generic' healthcare cost incurred by people in remission that is unrelated to treatment of depression. This cost was uplifted to 2016 prices using the HCHS index, resulting in a 2016 cost figure of £493 per year.

The figure of £493 was used to represent the cost of people in remission in the economic model. In the first year of remission following relapse, the annual cost of maintenance drug treatment incurred by people in remission was added to this figure, as well as the cost of 3 GP visits.

An overview of the healthcare costs associated with each health state in the guideline economic model and the methods for their estimation is provided in Table 278 and Table 279.

In the first 2 years of the model, the intervention cost of maintenance treatment was added onto the cost of the remission state, unless people relapsed within this period; in this case the intervention cost of maintenance treatment was added onto the cost of the remission state up to the point of relapse.

Update 2017

Table 278: Annual healthcare costs associated with the state of relapse in the guideline economic analysis (2016 prices)

Resource use element	Annual cost of relapse		Comments
	People remaining in relapse state in the next model cycle	Last year of relapse prior to moving to remission in the next model cycle	
Antidepressants	£77	£33	Cost reported in Byford et al. (2011) for non-remitters and remitters, respectively, multiplied by the estimated net ingredient cost per antidepressant prescription item ratio for 2015:2006 (Table 277). Cost for non-remitters was used in both calculations to reflect antidepressant usage over 12 months, as remitters in the study ceased pharmacological treatment within a period of less than 12 months, which is inconsistent with current recommended clinical practice for maintenance antidepressant treatment.
Concomitant medication	£102	£43	Cost reported in Byford et al. (2011) for non-remitters and remitters, respectively, multiplied by the estimated net ingredient cost per prescription item ratio for 2015:2006 (Table 277), weighted according to the concomitant medication usage reported in the study.
GP visits	£624	£464	Estimated by multiplying relevant resource use for non-remitters and remitters reported in Byford and colleagues (2011) with the GP unit cost of £36 per patient contact lasting 9.22 minutes for 2016 (Curtis and Burns, 2016).
GP phone calls	£150	£69	Estimated by multiplying resource use for non-remitters and remitters reported in Byford and colleagues (2011) with the GP unit cost of £28 per telephone consultation lasting 7.1 minutes (Curtis and Burns, 2016).
Psychological therapy contacts	£124	£124	Estimated by combining the percentage (16.8%) of people referred to and receiving IAPT psychological therapies in 2016 (NHS England, 2016) with the estimated weighted mean cost per course of IAPT treatment per person (£740), including people who completed treatment, those who dropped out, people who declined treatment and also people who were judged not to be suitable for treatment (Radhakrishnan et al., 2013), expressed in 2016 prices using the HCHS inflation index (Curtis and Burns, 2016). This cost was added to the annual cost of both people who remained in the relapse state and those who transitioned to the remission state in the next model cycle.
Psychiatrist contacts	£8	£5	Estimated by multiplying relevant resource use for non-remitters and remitters reported in Byford and colleagues (2011) with the 2016 NHS reference unit cost per contact with a mental health specialist team for adults and elderly of £121 (Department of Health, 2016).

Update 2017

Resource use element	Annual cost of relapse		Comments
	People remaining in relapse state in the next model cycle	Last year of relapse prior to moving to remission in the next model cycle	
Other specialist contacts	£90	£73	Estimated by multiplying relevant resource use for non-remitters and remitters reported in Byford and colleagues (2011) with the mean 2016 NHS reference unit cost per contact with outpatient services of £117 (Department of Health, 2016).
Hospitalisations [admissions]	£300	£263	Estimated by multiplying relevant resource use for non-remitters and remitters reported in Byford and colleagues (2011) with the mean 2016 NHS reference unit cost per admission in NHS care of £3,750 (Department of Health, 2016).
Accident and emergency attendances	£7	£6	Estimated by multiplying relevant resource use for non-remitters and remitters reported in Byford and colleagues (2011) with the mean 2015 NHS reference unit cost for accident and emergency services (outpatient attendances) of £147 (Department of Health, 2016).
TOTAL COST	**£1,483**	**£1,079**	

Table 279: Annual healthcare costs associated with the state of remission in the guideline economic analysis (2016 prices)

Resource use element	Annual cost of remission	Comments
Healthcare cost – all years of remission	**£493**	3-month healthcare cost of people having achieved remission obtained from graphs published by Byford and colleagues (2011), read using digital software (http://www.digitizeit.de), extrapolated to 12 months and uplifted to 2016 prices using the HCHS inflation index (Curtis and Burns, 2016).
Maintenance antidepressant therapy – 1st year extra cost	**£141**	Additional cost reflecting optimal duration of maintenance antidepressant therapy following remission, comprising of an annual antidepressant drug cost equal to that estimated for remitters and 3 GP contacts at the GP unit cost of £36 per patient contact lasting 9.22 minutes for 2016 (Curtis and Burns, 2016).

Update 2017

13.2.13 Cost of management of common side effects from antidepressant treatment

People who experienced common side effects were assumed to have one extra GP contact every 3 months costing £36 (Curtis & Burns, 2016) and to consume a cost of £10 per year for medication relating to the management of common side effects (e.g. paracetamol for the management of headaches).

13.2.14 Discounting

Costs and benefits were discounted at an annual rate of 3.5% as recommended by NICE (2014).

13.2.15 Handling uncertainty

Model input parameters were synthesised in a probabilistic analysis. This means that the input parameters were assigned probabilistic distributions (rather than being expressed as point estimates); this approach allowed more comprehensive consideration of the uncertainty characterising the input parameters and captured the non-linearity characterising the economic model structure. Subsequently, 10,000 iterations were performed, each drawing random values out of the distributions fitted onto the model input parameters. Results (mean costs and QALYs for each intervention) were averaged across the 10,000 iterations. This exercise provides more accurate estimates than those derived from a deterministic analysis (which utilises the mean value of each input parameter ignoring any uncertainty around the mean), by capturing the non-linearity characterising the economic model structure (Briggs et al., 2006).

The distributions of the hazard ratios of all treatments versus pill placebo (reflecting clinical management) were obtained from the NMAs, defined directly from values recorded in each of the 10,000 iterations performed in WinBUGS. The distributions of risk ratios of antidepressants versus placebo that were utilised in analyses in people at medium risk of relapse were assigned a log-normal distribution.

The baseline risk of relapse after a single (first) episode and the risk of recovery were both determined by a Weibull distribution, as described earlier in methods. The probability distributions of the Weibull parameters (gamma and lambda) were defined directly from values recorded in each of the 10,000 iterations performed in WinBUGS. This allowed the correlation between the Weibull parameters to be taken into account. The hazard ratio of the risk of relapse for every additional depressive episode was given a log-normal distribution.

Utility values were assigned a beta distribution after applying the method of moments on data reported in the relevant literature. The proportion of women in the sample and the proportion of people experiencing side effects were also assigned a beta distribution. The risk ratio of mortality was assigned a log-normal distribution.

Uncertainty in intervention costs was taken into account by assigning probability distributions to the number of GP contacts and the number of individually delivered psychological therapy sessions. The number of therapist sessions per person attending group psychological interventions was not assigned a probability distribution because the number of group sessions remains the same, whether a participant attends the full course of treatment or a lower number of sessions. Drug acquisition costs were not given a probability distribution as these costs are set and are characterised by minimal uncertainty. However, if people receiving maintenance pharmacological therapy attended fewer GP visits than the mode in the second year of maintenance treatment, then they were assumed to be prescribed smaller amounts of medication than optimal, and to subsequently incur lower drug acquisition costs. Unit costs of healthcare staff (GPs and clinical psychologists) were assigned a normal

Update 2017

distribution. Healthcare costs associated with the states of relapse and recovery were assigned a gamma distribution.

Table 280 provides details on the types of distributions assigned to each input parameter and the methods employed to define their range.

Table 280: Input parameters (deterministic values and probability distributions) that informed the economic models of interventions for relapse prevention in adults with depression that is in remission

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
General characteristics of population			Kessler et al., 2005; Fernandez-Pujals et al., 2015; GC expert advice
Age of onset (years)	32	No distribution	
Mean interval between episodes (years)	2	No distribution	GC expert advice
Number of previous episodes	1	No distribution	GP expert advice
- medium risk of relapse	3	No distribution	McManus et al., 2016; weighted prevalence of depression 2.9% in men, 3.7% in women, survey sample N=7,546
- high risk of relapse	0.56	Beta: α=279; β=219	
Proportion of women			
Risk ratios vs pill placebo – people at medium risk of relapse who remitted following acute pharmacological treatment			
		Log-normal:	Guideline pairwise meta-analysis
Citalopram (SSRI)	0.61	95% CI 0.56 to 0.68	
Venlafaxine (SNRI)	0.66	95% CI 0.55 to 0.78	
Amitriptyline (TCA)	0.70	95% CI 0.43 to 1.14	
Mirtazapine	0.67	95% CI 0.45 to 0.98	
Hazard ratios vs pill placebo – people at high risk of relapse who remitted following acute pharmacological treatment			
Antidepressant	0.508	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
MBCT (antidepressant tapering)	0.446		
MBCT and antidepressant	0.346		
Hazard ratios vs pill placebo – people at high risk of relapse who remitted following acute pharmacological treatment: sensitivity analysis			
Antidepressant	0.509	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
MBCT (antidepressant tapering)	0.450		
MBCT and antidepressant	0.346		
Group CT and antidepressant	0.393		
Hazard ratios vs pill placebo – people at medium or high risk of relapse who remitted following acute psychological treatment			
CT	0.717	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
Fluoxetine	0.965		
No treatment (wait list)	1.335		
MBCT (sensitivity analysis only)	0.910		

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
group CT (sensitivity analysis only)	1.016		
Hazard ratios vs pill placebo – people at high risk of relapse who remitted following acute combination treatment			
Combination therapy Antidepressant Psychological therapy (antidepressant tapering)	0.337 0.425 0.703	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
Baseline risk of relapse after a single (first) episode Weibull distribution – lambda Weibull distribution – gamma Hazard ratio – new vs previous episode	 0.095 0.611 1.15	 WinBUGS output WinBUGS output Log-normal: 95% CI 1.11 to 1.18	Synthesis of data from Eaton et al., 2008 and Mattisson et al., 2007, using a Bayesian approach – fixed effects model Kessing & Andersen, 1999
Risk of recovery Weibull distribution – lambda Weibull distribution – gamma	 1.171 0.440	 WinBUGS output WinBUGS output	Synthesis of data from Gonzales et al., 1985; Holma et al., 2008; Keller & Shapiro, 1981; Keller et al., 1984 & 1992; Mueller et al., 1996; Skodol et al., 2011 & Stegenga et al., 2012, using a Bayesian approach – random effects model
Probability of developing common side effects – SSRIs alone or in combination – SNRIs – TCAs – mirtazapine	 0.117 0.150 0.152 0.163	 Beta: α=2,752; β=20,868 Beta: α=714; β=4,048 Beta: α=118; β=658 Beta: α=147; β=754	Anderson et al., 2012
Mortality Risk ratio – depressed vs non-depressed Baseline mortality – non-depressed	 1.52 Age/sex spec	 Log-normal: 95% CI 1.45 to 1.59 No distribution	 Cuijpers et al., 2014 Mortality statistics for the UK population (ONS, 2015)
Utility values Less severe depression More severe depression Remission/recovery Disutility due to side effects	 0.60 0.42 0.81 0.09	 Beta: α=182; β=122 Beta: α=54; β=75 Beta: α=531; β=125 Beta: α=6; β=59	Distributions determined using method of moments, based on data reported in Sobocki et al., 2006 & 2007, Sullivan et al., 2004 and further assumptions
Intervention costs – resource use Number of GP visits – drug treatment			Probabilities assigned to numbers of sessions

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
1st year	6	0.70: 6, 0.20: 4-5, 0.10: 2-3	Number of visits based on GC expert opinion; probabilities based on assumption. If number of GP visits in 2nd year of pharmacological treatment was lower than 3, only 50% of the drug acquisition cost was incurred and 50% of annual GP contacts due to side effects were made
2nd year	3	0.70: 3, 0.30: 1-2	
tapering	1	0.70: 1, 0.30: 2	
Number of GP visits – clinical management (pill placebo)			See note on GP visits in 2nd year of maintenance drug treatment
1st year	3	0.70: 3, 0.20: 1-2, 0.10: 0	Participants missing one or more group sessions assumed not to be replaced by others; therefore no impact on total intervention cost
2nd year	1	0.70: 1, 0.30: 0	
Number of GP visits - side effects (annual)	4	2 or 4 in second year	
Number of GP visits – psychol. therapy	2	0.60: 2, 0.40: 1	
Number of MBCT group sessions	12	No distribution	Number of visits based on GC expert opinion; probabilities based on assumption
Number of group CT sessions	8	No distribution	
Number of CT/CBT individual sessions	10	0.60: 10, 0.20: 8-9, 0.15: 6-7, 0.05: 1-5	
Intervention costs - unit costs			
Drug acquisition costs	See Table 272	No distribution	National drug tariff, January 2017
GP unit cost	£36	Normal, SE=0.05*mean	Curtis & Burns, 2016; distribution based on assumption
Clinical psychologist unit cost	£97	Normal, SE=0.05*mean	See Table 274; distribution based on assumption
Annual NHS health state cost		Gamma	Based primarily on cost data reported in Byford et al., 2011, supplemented by data from Curtis & Burns, 2016; NHS England, 2016; and Radhakrishnan et al., 2013, expressed in 2016 prices using the HCHS inflation index (Curtis & Burns, 2016). For more details see Table 278 and Table 279; distribution based on assumption
Relapse - remaining in state	£1,483	SE=0.20*mean	
Relapse - final year before remission	£1,079		
Remission	£493		
Remission – 1st year extra cost	£141		
Annual discount rate	0.035	No distribution	Applied to both costs and outcomes. NICE, 2014

Update 2017

A number of deterministic one- and n- way (combined) sensitivity analyses were undertaken to explore the impact of alternative hypotheses on the results. The following scenarios were explored alone or in combination:

- Change (increase) in the number of previous episodes, resulting in an increase in the risk of relapse; the number of previous episodes was increased from 1 to 2 in people at medium risk of relapse and from 3 to 5 in people at high risk of relapse
- Change in the severity of previous episodes, as reflected in respective health state utility values for less severe depression and more severe depression; under this scenario, people at medium risk of relapse were assumed to experience more severe depression if they relapsed and people at high risk of relapse were assumed to experience less severe depression if they relapsed.
- Use of utility values for less severe depression and more severe depression reported in Mann and colleagues (2009)
- Setting the cost of GP visits associated with clinical management (pill placebo) at zero
- Change in the cost associated with the state of relapse by ± 50%
- Reduction in the number of individual CBT/CT sessions down to 4 (from 10, which was the number used in base-case analyses), to reflect more closely routine UK clinical practice
- Assuming a shorter relapse preventive effect of psychological interventions, by applying the hazard ratios of psychological interventions onto the baseline risk of relapse over the first year of the economic analysis only (and not in the first and second year, as in the base-case analysis). Under this scenario, the relapse preventive effect of combination therapies in the second year of the economic analysis was assumed to equal the effect of their pharmacological intervention component. This scenario was explored because the evidence on the long term effects of psychological interventions in relapse prevention (i.e. beyond one year and closer to two years) is limited and existing evidence suggests a reduction in this effect (Kuyken 2015).

Update 2017

13.2.16 Presentation of the results

Results of the economic analysis are presented as follows:

Results are reported separately for each cohort examined in the economic model. In each analysis, mean total costs and QALYs are presented for each intervention, averaged across 10,000 iterations of the model. An incremental analysis is provided for each cohort, in table format, where all options have been listed 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) or by extended dominance (that is, they are less effective and more costly than a linear combination of two alternative 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 (QALY) 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/QALY (NICE 2008, Social value judgements) is the most cost-effective option.

In addition to ICERs, the mean net monetary benefit (NMB) of each intervention is presented. This is defined by the following formula:

$$NMB = E \bullet \lambda - C$$

where E and C are the effectiveness (number of QALYs) and costs associated with the treatment option, respectively, and λ is the level of the willingness-to-pay (WTP) per unit of effectiveness, set at the NICE lower cost effectiveness threshold of £20,000/QALY (NICE, 2008). The intervention with the highest NMB is the most cost-effective option (Fenwick et al., 2001).

Incremental mean costs and effects (QALYs) of each maintenance intervention versus clinical management (with antidepressant drug tapering if relevant) are also presented in the form of cost effectiveness planes.

The probability of each intervention being the most cost-effective option at the NICE lower cost effectiveness threshold of £20,000/QALY is also provided, calculated as the proportion of iterations (out of the 10,000 iterations run) in which the intervention has had the highest NMB among all interventions considered in the analysis. These probabilities are also summarised in cost-effectiveness acceptability curves (CEACs), which show the probability of each intervention being cost-effective at various cost-effectiveness thresholds.

Finally, the mean ranking in terms of cost effectiveness is provided for each intervention (out of the 10,000 iterations run), with higher rankings suggesting higher cost effectiveness.

13.2.17 Validation of the economic model

The economic model (including the conceptual model and the identification and selection of input parameters) was developed by the health economist in collaboration with a health economics sub-group formed by members of the Guideline Committee. As part of the model validation, all inputs and model formulae were systematically checked; 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; moreover, a number of parameters, such as efficacy (risk and odds ratios), intervention costs, and number of previous episodes (which differ between populations at medium and high risk of relapse) were set at the same value across interventions and analyses, to explore whether total costs and benefits across interventions and analyses became equal, as expected. The base-case results and results of sensitivity analyses were discussed with the Guideline Committee to confirm their plausibility. In addition, the economic model (excel spreadsheet) and this chapter were checked for their validity and accuracy by a health economist that was external to the guideline development team.

Update 2017

13.3 Results of the economic analysis

13.3.1 People at medium risk of relapse who remitted following acute pharmacological treatment

The base-case results of the analysis are presented in Table 281. Maintenance treatment with SSRIs, SNRIs, TCAs or mirtazapine was less cost-effective than clinical management and drug tapering in people at medium risk of relapse who remitted following acute pharmacological treatment with SSRIs, SNRIs, TCAs or mirtazapine, respectively and who were assumed to experience less severe depression if they relapsed. Maintenance treatment with SSRIs resulted in slightly higher benefits (QALYs) at an additional cost of £293,305/QALY, which is well above the NICE cost-effectiveness threshold of £20,000-£30,000/QALY. Maintenance treatment with SNRIs, TCAs and mirtazapine was dominated by clinical management and drug tapering (i.e. it resulted in fewer QALYs and higher costs compared with clinical management). Results of deterministic analysis were similar.

Table 281: Results of economic modelling: interventions for people at medium risk of relapse who remitted following acute pharmacological treatment and who experienced less severe depression if they relapsed (mean values from probabilistic analysis)

Maintenance treatment option	Mean /person		ICER (£/QALY)	NMB (£) /person	Prob best[1]	Mean ranking
	QALY	Cost				
People who remitted following acute SSRI treatment						
SSRI	6.837	5,055	293,305	131,689	0.30	1.70
Clinical management (SSRI tapering)	6.837	4,944		131,793	0.70	1.30
People who remitted following acute SNRI treatment						
Clinical management (SNRI tapering)	6.837	4,944	Dominant	131,792	0.93	1.07
SNRI	6.829	5,104		131,482	0.07	1.93
People who remitted following acute TCA treatment						
Clinical management (TCA tapering)	6.837	4,944	Dominant	131,792	0.91	1.09
TCA	6.826	5,121		131,405	0.09	1.91
People who remitted following acute mirtazapine treatment						
Clinical management (Mirt tapering)	6.837	4,944	Dominant	131,792	0.91	1.09
Mirtazapine	6.826	5,095		131,430	0.09	1.91

1 at the NICE lower cost-effectiveness threshold of £20,000/QALY

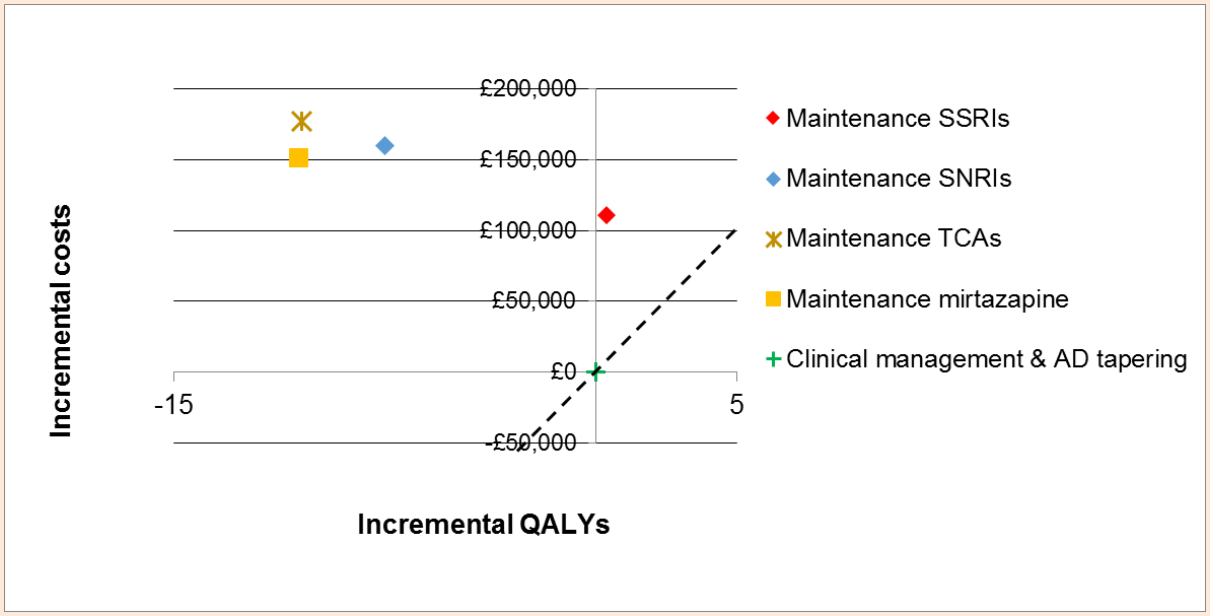
Prob: probability; SNRI: serotonin–norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor; TCA: tricyclic antidepressant

Figure 26 provides the cost effectiveness plane of the analysis. Each intervention is placed on the plane according to its incremental costs and QALYs compared with clinical management and antidepressant drug tapering, which is placed at the origin. The slope of the dotted line indicates the NICE lower cost effectiveness threshold, suggesting that maintenance pharmacological treatment is not cost-effective compared with clinical management and antidepressant drug tapering for people at medium risk of relapse who remitted following acute pharmacological treatment (since all maintenance pharmacological treatments lie on the left side of the dotted line). It is noted that results for each maintenance pharmacological intervention versus clinical management and drug tapering refer to different study populations, depending on the acute pharmacological treatments they received, and therefore estimating the relative cost effectiveness between different maintenance pharmacological treatments is not relevant or appropriate.

Figure 26 Cost effectiveness plane of maintenance pharmacological interventions for people at medium risk of relapse who remitted following acute pharmacological treatment and who experienced less severe depression if

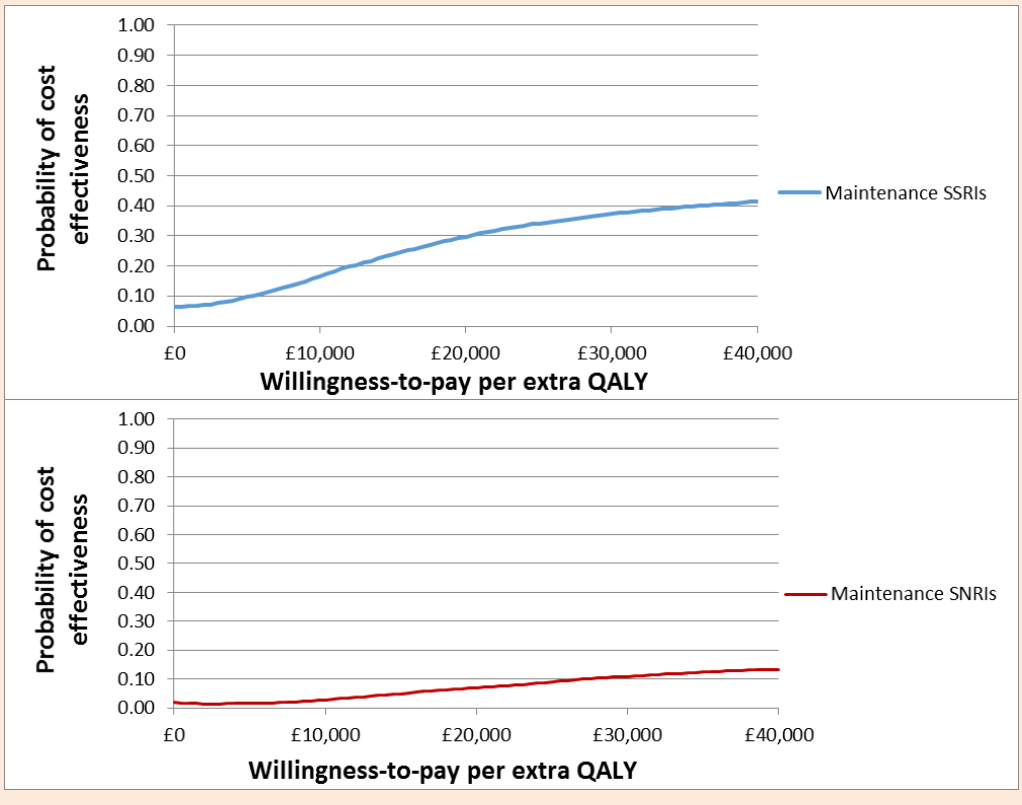
Update 2017

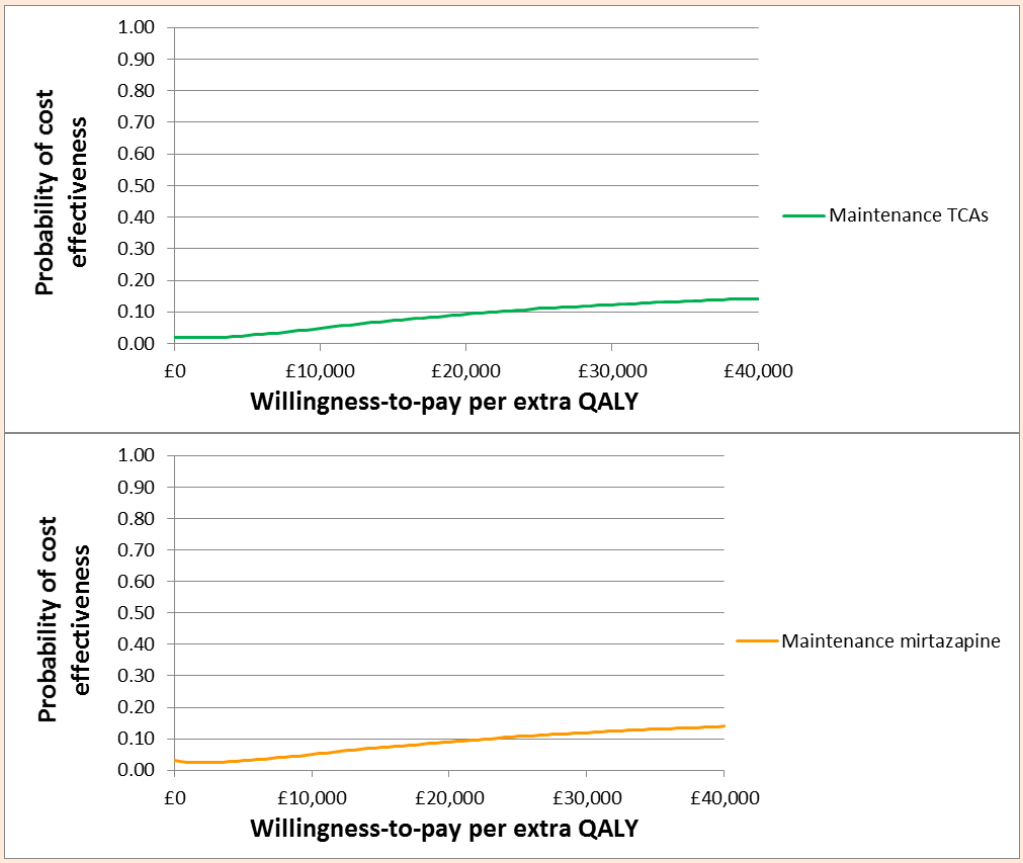
they relapsed – incremental costs and QALYs versus clinical management and antidepressant drug tapering per 1,000 adults



The probability of each pharmacological intervention being cost-effective compared with clinical management and drug tapering was very low and ranged from 0.07 for SNRIs to 0.30 for SSRIs at the NICE lower cost-effectiveness threshold of £20,000/QALY. The probability of each intervention being cost-effective compared with clinical management and drug tapering at various levels of WTP per QALY gained (i.e. at a range of cost effectiveness thresholds) is shown in Figure 27.

Figure 27. Cost-effectiveness acceptability curves of interventions for people at medium risk of relapse who remitted following acute pharmacological treatment and who experienced less severe depression if they relapsed





In deterministic sensitivity analysis, increasing the number of previous episodes from 1 to 2 had no impact on the results and conclusions of the analysis.

Assuming that future relapse episodes were more severe in terms of the associated utility value resulted in maintenance treatment with SSRIs becoming cost-effective, with an ICER versus clinical management and SSRI tapering of £7,451/QALY. Maintenance treatment with SNRIs, TCAs and mirtazapine became more effective than clinical management and drug tapering, but the resulting ICERs were above the NICE lower cost effectiveness threshold (ranging from £26,079/QALY for SNRIs to £66,334/QALY for TCAs).

Combining the two scenarios, i.e. assuming that people had 2 previous episodes and relapsed to more severe depression resulted in SSRIs, SNRIs and mirtazapine becoming more cost-effective than clinical management and drug tapering, with ICERs of £4,963/QALY, £14,210/QALY and £18,167/QALY, respectively. The ICER of maintenance treatment with TCAs versus clinical management and TCA tapering was still above the NICE lower cost effectiveness threshold, at £25,349/QALY.

Use of a higher utility value for less severe depression from a different source (which reduced the scope for QALY improvements following relapse prevention) or assuming a zero intervention cost for clinical management (so that it reflected no treatment in terms of cost) further reduced the cost effectiveness of pharmacological maintenance treatment compared with clinical management in this population.

Changing the cost of the relapse health state by 50% had no impact on the results and conclusions of the analysis.

Update 2017

13.3.2 People at high risk of relapse who remitted following acute pharmacological treatment

13.3.2.1 Base-case analysis

The base-case results of the analysis are presented in Table 282. The most cost-effective maintenance treatment option for people at high risk of relapse who remitted following acute pharmacological treatment and who were assumed to relapse to more severe depression was MBCT combined with clinical management (antidepressant tapering), with an ICER versus clinical management (antidepressant tapering) alone of £1,013/QALY. The options of MBCT combined with antidepressant treatment and of maintenance antidepressant treatment alone were dominated by absolute or extended dominance, respectively. The probability of MBCT and antidepressant tapering being cost-effective was 0.83 at the NICE lower cost-effectiveness threshold of £20,000/QALY. MBCT combined with maintenance antidepressant treatment was the second most cost-effective option, followed by maintenance antidepressant treatment alone, which was more cost-effective (third best option) than clinical management (antidepressant tapering). Results of base-case deterministic analysis were very similar.

Table 282: Results of economic modelling: interventions for people at high risk of relapse who remitted following acute pharmacological treatment and who experienced more severe depression if they relapsed – base-case analysis (mean values from probabilistic analysis)

Maintenance treatment option	Mean, /person		ICER (£/QALY)	NMB (£) /person	Prob best[1]	Mean ranking
	QALY	Cost				
MBCT & clinical management (AD tapering)	6.737	5,177	1,013	129,564	0.83	1.18
MBCT & AD	6.731	5,300	Dominated	129,310	0.16	1.96
AD	6.711	5,163	Ext Dom	129,052	0.01	2.87
Clinical management (AD tapering)	6.673	5,112		128,344	0.00	3.99

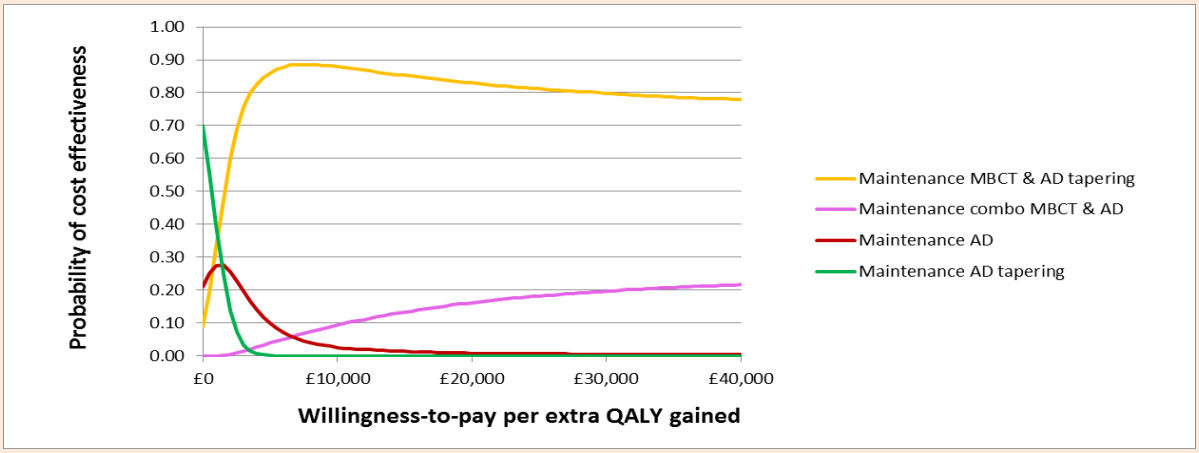
Notes:
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
AD: antidepressant; Ext Dom: extendedly dominated; MBCT: mindfulness-based cognitive therapy; Prob: probability

The probability of each intervention being cost-effective at various levels of WTP per QALY gained is shown in Figure 28.

Figure 28. Cost-effectiveness acceptability curves of interventions for people at high risk of relapse who remitted following acute pharmacological treatment and

Update 2017

who experienced more severe depression if they relapsed – base-case analysis



Conclusions and rankings of interventions in terms of cost effectiveness remained the same under the vast majority of scenarios explored in deterministic sensitivity analysis, including:

- increase in the number of previous episodes from 3 to 5
- future relapse episodes being assumed to be less severe in terms of the associated utility value
- use of a higher utility value for more severe depression from a different source (which reduced the scope for QALY improvements following relapse prevention)
- use of a zero intervention cost for clinical management (so that it reflected no treatment in terms of cost)
- change in the cost of relapse by ±50%

Assuming that the preventive effect of MBCT lasted only one year resulted in the combination of MBCT plus antidepressant treatment becoming the most cost-effective intervention, followed by MBCT plus clinical management (antidepressant tapering), then antidepressant treatment alone, and, finally, clinical management and antidepressant drug tapering.

Update 2017

13.3.2.2 Sensitivity analysis including additional interventions

The additional intervention included in this sensitivity analysis was group CT added onto maintenance antidepressant treatment. Results are provided in Table 283. Results remained unchanged in terms of the most cost-effective interventions and the relative ranking of the interventions included in the base-case analysis. Group CT added onto maintenance antidepressant treatment became the third most cost-effective option; MBCT with clinical management (antidepressant tapering) remained the most effective and cost-effective option, with a 0.76 probability of being the most cost-effective option at the NICE lower cost-effectiveness threshold of £20,000/QALY. Results of deterministic analysis were very similar.

Table 283: Results of economic modelling: interventions for people at high risk of relapse who remitted following acute pharmacological treatment and who experienced more severe depression if they relapsed – sensitivity analysis (mean values from probabilistic analysis)

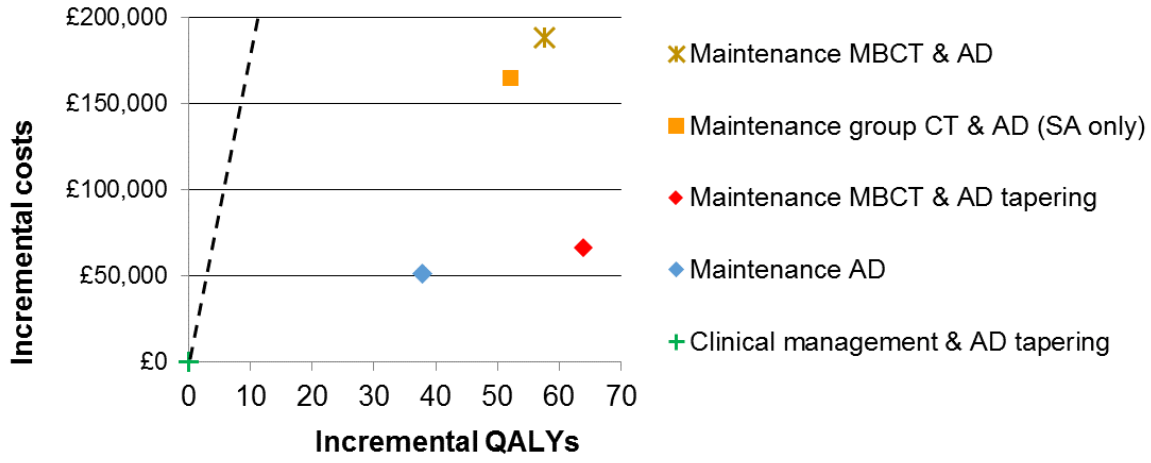
Maintenance treatment option	Mean /person		ICER (£/QALY)	NMB (£) /person	Prob best[1]	Mean ranking
	QALY	Cost				
MBCT & clinical management (AD tapering)	6.737	5,178	1,038	129,554	0.76	1.33
MBCT & AD	6.730	5,300	Dominated	129,309	0.09	2.28

Maintenance treatment option	Mean /person		ICER (£/QALY)	NMB (£) /person	Prob best[1]	Mean ranking
	QALY	Cost				
Group CT & AD	6.725	5,276	Dominated	129,220	0.15	2.68
AD	6.711	5,163	Ext Dom	129,050	0.00	3.73
Clinical management (AD tapering)	6.673	5,112		128,344	0.00	4.98

Notes:
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
AD: antidepressant; CT: cognitive therapy; Ext Dom: extendedly dominated; MBCT: mindfulness-based cognitive therapy; Prob: probability

Figure 29 provides the cost effectiveness plane of the base-case analysis, including the additional intervention assessed in sensitivity analysis. Each intervention is placed on the plane according to its incremental costs and QALYs compared with clinical management and antidepressant drug tapering, which is placed at the origin. The slope of the dotted line indicates the NICE lower cost effectiveness threshold, suggesting that all maintenance pharmacological treatments assessed in the analysis are cost-effective compared with clinical management and antidepressant drug tapering for people at high risk of relapse who remitted following acute pharmacological treatment (since all maintenance pharmacological treatments lie on the right side of the dotted line).

Figure 29 Cost effectiveness plane of maintenance interventions for people at high risk of relapse who remitted following acute pharmacological treatment and who experienced more severe depression if they relapsed – incremental costs and QALYs versus clinical management and antidepressant drug tapering per 1,000 adults. Base-case and sensitivity analysis

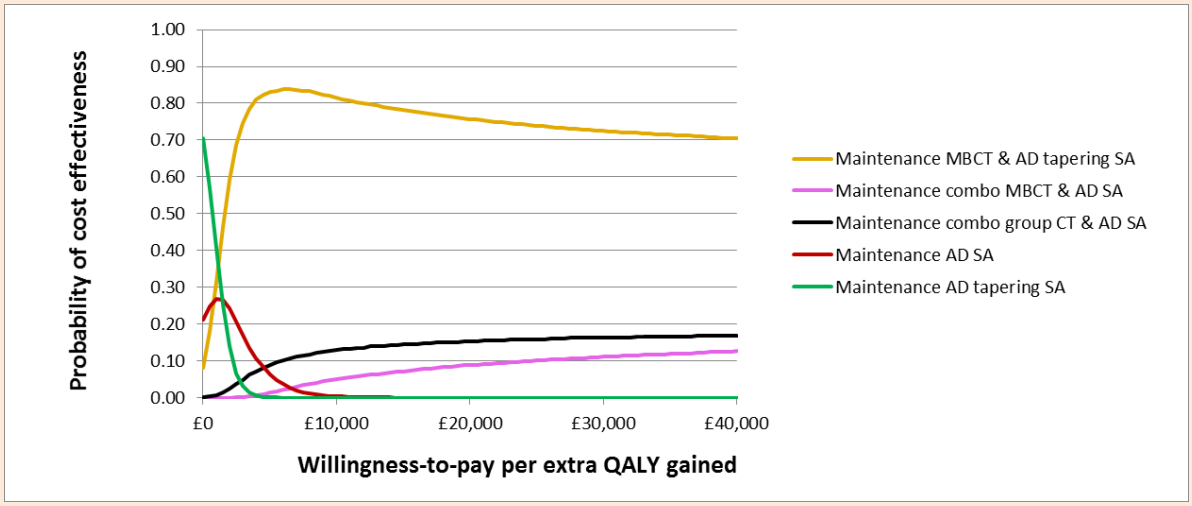


The probability of each intervention being cost-effective in the sensitivity analysis that included group CT at various levels of WTP per QALY gained is shown in Figure 30.

Figure 30 Cost-effectiveness acceptability curves of interventions for people at high risk of relapse who remitted following acute pharmacological treatment and

Update 2017

who experienced more severe depression if they relapsed – sensitivity analysis



13.3.34 People at medium risk of relapse who remitted following acute psychological treatment

The base-case results of this analysis are presented in Table 284. The most cost-effective maintenance treatment option for people at medium risk of relapse to less severe depression who remitted following acute psychological treatment (CT) was clinical management, followed by no treatment. Maintenance CT was the most effective option but also the one with the highest cost, with an ICER of £51,985/QALY versus clinical management; it was the third most cost-effective option, above fluoxetine which was the least cost-effective option. The probability of clinical management being the most cost-effective option was 0.58 at the NICE lower cost-effectiveness threshold of £20,000/QALY. The relative cost effectiveness between interventions was the same in deterministic analysis.

Update 2017

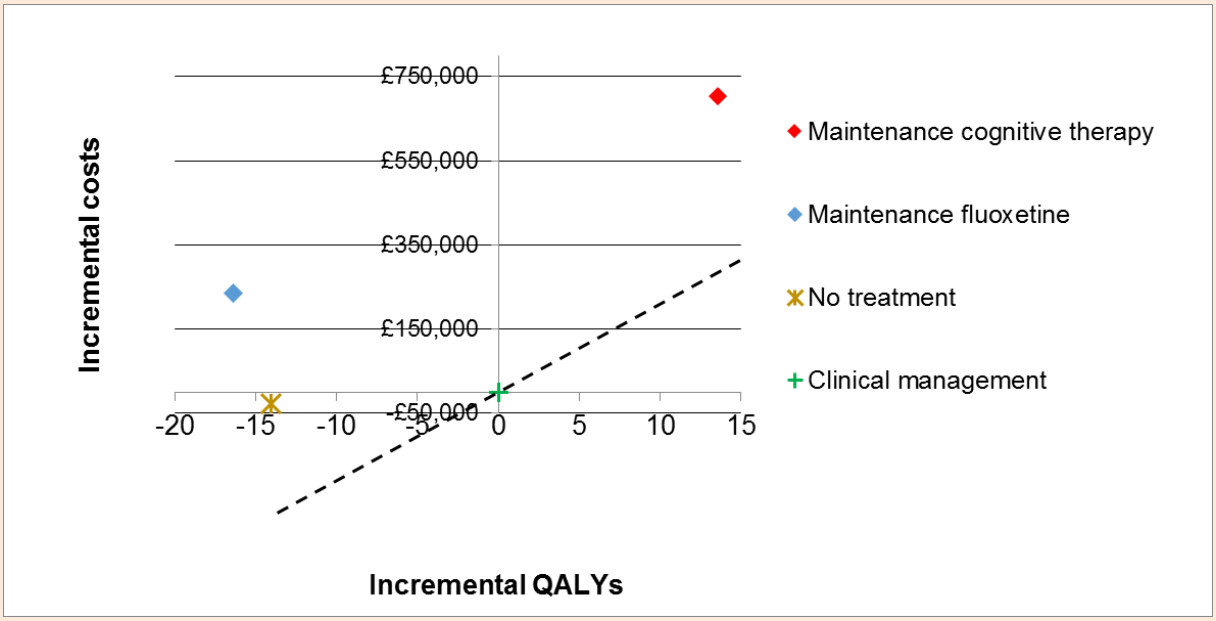
Table 284: Results of economic modelling: interventions for people at medium risk of relapse who remitted following acute psychological treatment and who experienced less severe depression if they relapsed (mean values from probabilistic analysis)

Maintenance treatment option	Mean /person		ICER (£/QALY)	NMB (£) /person	Prob best[1]	Mean ranking
	QALY	Cost				
CT	6.850	5,602	51,985	131,405	0.04	2.98
Clinical management	6.837	4,899	2,007	131,837	0.58	1.46
No treatment (wait list)	6.823	4,871		131,584	0.37	2.20
Fluoxetine	6.820	5,134	Dominated	131,275	0.01	3.36

Notes:
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
CT: cognitive therapy; Prob: probability

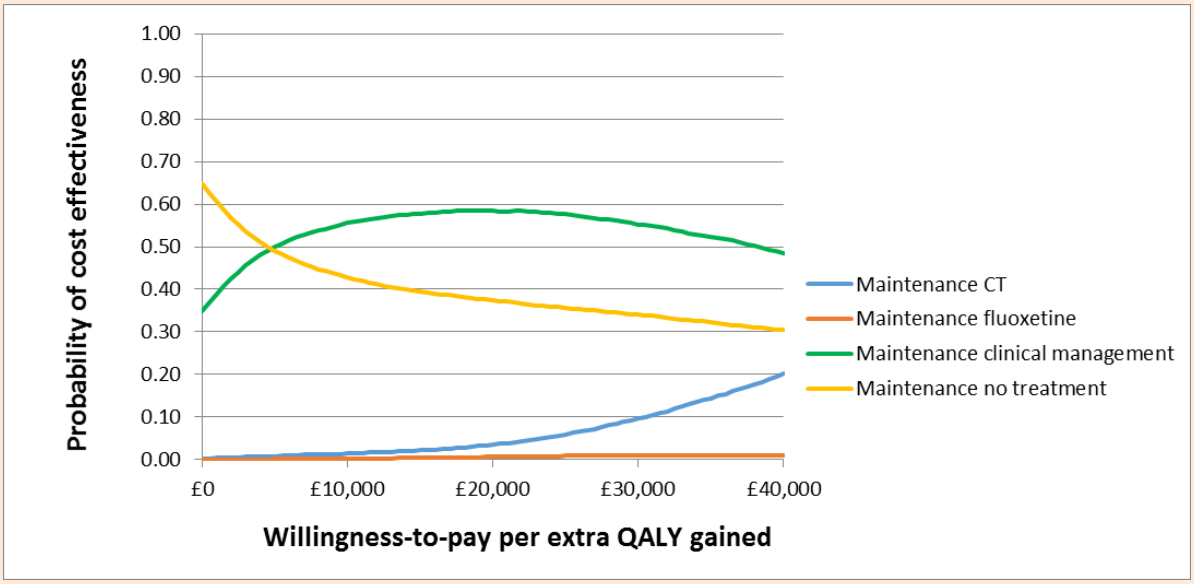
Figure 31 provides the cost effectiveness plane of the analysis. Each intervention is placed on the plane according to its incremental costs and QALYs compared with clinical management, which is placed at the origin. The slope of the dotted line indicates the NICE lower cost effectiveness threshold, suggesting that maintenance treatments and no treatment are not cost-effective compared with clinical management for people at medium risk of relapse who remitted following acute psychological treatment (since all options lie on the left side of the dotted line).

Figure 31 Cost effectiveness plane of maintenance treatments (or no treatment) for people at medium risk of relapse who remitted following acute psychological treatment and who experienced less severe depression if they relapsed – incremental costs and QALYs and antidepressant drug tapering per 1,000 adults



The probability of each intervention being cost-effective in people at medium risk of relapse who remitted following acute psychological treatment and who experienced less severe depression if they relapsed at various levels of WTP per QALY gained is shown in Figure 32.

Figure 32 Cost-effectiveness acceptability curves of interventions for people at medium risk of relapse who remitted following acute psychological treatment and who experienced less severe depression if they relapsed



In deterministic sensitivity analysis, increasing the number of previous depressive episodes (and therefore the risk of future relapses) from 1 to 2 did not have any impact on the conclusions of the analysis and the ranking of interventions.

Update 2017

Assuming that future relapse episodes were more severe in terms of the associated utility value resulted in maintenance CT becoming the second most cost-effective option, above no treatment.

Use of a higher utility value for less severe depression from an alternative source resulted in maintenance CT becoming the least cost-effective option.

Assuming a zero intervention cost for clinical management (so that it reflected no treatment in terms of cost) further improved the cost effectiveness of this option, as expected.

Assuming a 50% reduction in the cost of the relapse state resulted in maintenance CT becoming the least cost-effective option. A 50% increase in the cost of relapse did not have any impact on the results.

Reducing the number of sessions of CT to 4 had a significant impact on the results: CT became the most cost-effective intervention, with an ICER of £17,742/QALY versus clinical management. The relative ranking of the other interventions was not affected, as expected.

Assuming that the relapse preventive effect of CT lasted only one year resulted in CT becoming the least cost-effective option.

In a combined scenario where maintenance CT comprised 4 sessions and its preventive effect lasted only 1 year, CT was the second most cost-effective option following clinical management. If the assumption that people who relapse experience more severe depression was added onto this scenario, then CT became the most cost-effective option.

13.3.4 People at high risk of relapse who remitted following acute psychological treatment

Update 2017

13.3.4.1 Base-case analysis

The most cost-effective maintenance treatment option for people at high risk of relapse to more severe depression who remitted following acute psychological treatment (CT) was clinical management followed by maintenance CT, which was marginally less cost-effective (its ICER versus clinical management was £20,971/QALY). Third most cost-effective option was fluoxetine and, finally, no treatment (wait list) was the least cost-effective option in this population. Maintenance CT was the most effective but also the costliest treatment option. The probability of clinical management being the most cost-effective option was 0.39 at the NICE lower cost-effectiveness threshold of £20,000/QALY, indicating the uncertainty underlying the results. The base-case results of this analysis are shown in Table 285. Results of deterministic analyses were similar.

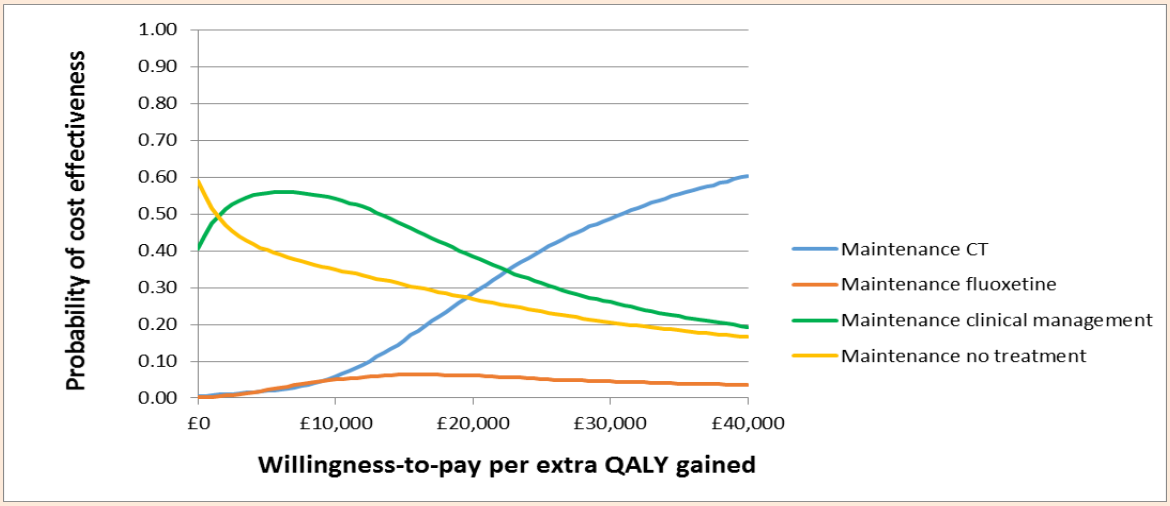
Table 285: Results of economic modelling: interventions for people at high risk of relapse who remitted following acute psychological treatment and who experienced more severe depression if they relapsed (mean values from probabilistic analysis)

Maintenance treatment option	Mean /person		ICER (£/QALY)	NMB (£) /person	Prob best[1]	Mean ranking
	QALY	Cost				
CT	6.705	5,742	20,971	128,357	0.28	2.05
Clinical management	6.673	5,068	270	128,389	0.39	2.00
Fluoxetine	6.660	5,293	dominated	127,897	0.06	3.14
No treatment (wait list)	6.641	5,059		127,759	0.27	2.81

Notes:
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
CT: cognitive therapy; Prob: probability

The probability of each intervention being cost-effective in people at high risk of relapse who remitted following acute psychological treatment and who experienced more severe depression if they relapsed at various levels of WTP per QALY gained is shown in Figure 33.

Figure 33. Cost-effectiveness acceptability curves of interventions for people at high risk of relapse who remitted following acute psychological treatment and who experienced more severe depression if they relapsed – base-case analysis



In deterministic sensitivity analysis, increasing the number of previous depressive episodes (and therefore the risk of future relapses) from 3 to 5 resulted in maintenance CT becoming the most cost-effective option.

Assuming that future relapse episodes had the utility of less severe depression (instead of more severe) resulted in no treatment becoming the second most cost-effective treatment option, below clinical management and above CT and fluoxetine (the latter was the least cost-effective option under this scenario).

Use of a higher utility value for more severe depression resulted in no treatment becoming the third most cost-effective option, following CT, with fluoxetine being ranked fourth.

Assuming a zero intervention cost for clinical management (so that it reflected no treatment in terms of cost) did not have any impact on the results of the analysis.

Applying a 50% change in the cost had no impact on the results either.

Reducing the number of sessions of CT to 4 further improved the cost effectiveness of CT, the ICER of which versus clinical management dropped at £6,794/QALY, thus becoming the most cost-effective option.

Assuming that the preventive effect of CT lasted only one year had no impact on the results. In the scenario where CT comprised 4 visits and its preventive effect lasted only one year, CT was the most cost-effective option with an ICER versus clinical management of £11,488/QALY.

13.3.4.28 Sensitivity analysis including additional interventions

The additional interventions included in this sensitivity analysis were MBCT and group CT. Results are provided in Table 286. MBCT was the most cost-effective option, with a 0.35 probability of being most cost-effective at the NICE lower cost-effectiveness threshold of £20,000/QALY. Clinical management was the second best option, followed by individual CT and then group CT. Fluoxetine was the fifth most cost-effective option and no treatment was

Update 2017

the least cost-effective among options assessed. In deterministic sensitivity analysis, group CT ranked in a higher position than individual CT (they were ranked third and fourth, respectively).

Table 286: Results of economic modelling: interventions for people at high risk of relapse who remitted following acute psychological treatment and who experienced more severe depression if they relapsed – sensitivity analysis (mean values from probabilistic analysis)

Maintenance treatment option	Mean /person		ICER (£/QALY)	NMB (£) /person	Prob best[1]	Mean ranking
	QALY	Cost				
CT	6.705	5,742	28,158	128,357	0.14	3.23
MBCT	6.685	5,168	8,555	128,523	0.35	2.36
group CT	6.674	5,162	Ext Dom	128,315	0.25	2.99
Clinical management	6.673	5,068	270	128,389	0.22	3.16
Fluoxetine	6.660	5,293	Dominated	127,897	0.04	4.60
No treatment (wait list)	6.641	5,059		127,759	0.00	4.65

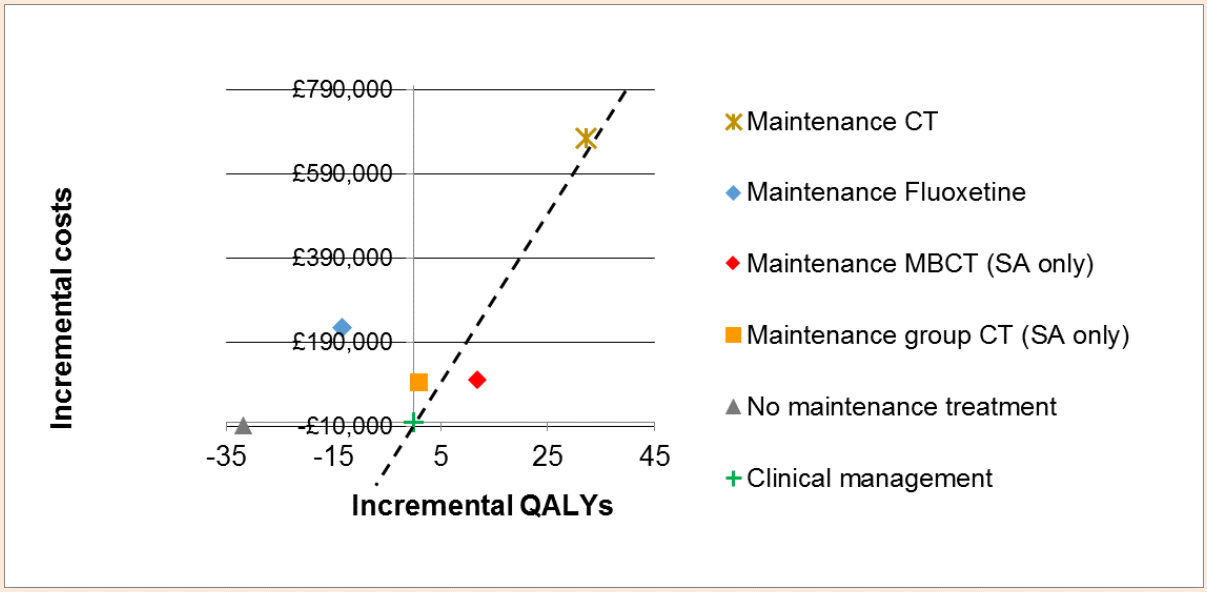
Notes:
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
CT: cognitive therapy; Ext Dom: extendedly dominated; MBCT: mindfulness-based cognitive therapy; Prob: probability

Figure 34 provides the cost effectiveness plane of the base-case analysis, including the additional interventions assessed in sensitivity analysis. Each intervention is placed on the plane according to its incremental costs and QALYs compared with clinical management and antidepressant drug tapering, which is placed at the origin. The slope of the dotted line indicates the NICE lower cost effectiveness threshold, suggesting that only MBCT is cost-effective compared with clinical management for people at high risk of relapse who remitted following acute psychological treatment (since this is the only maintenance intervention lying on the right side of the dotted line). Maintenance CT is marginally less cost-effective than clinical management, lying on the left of but very close to the dotted line of cost effectiveness.

Figure 34 Cost effectiveness plane of maintenance interventions for people at high risk of relapse who remitted following acute psychological treatment and who experienced more severe depression if they relapsed – incremental

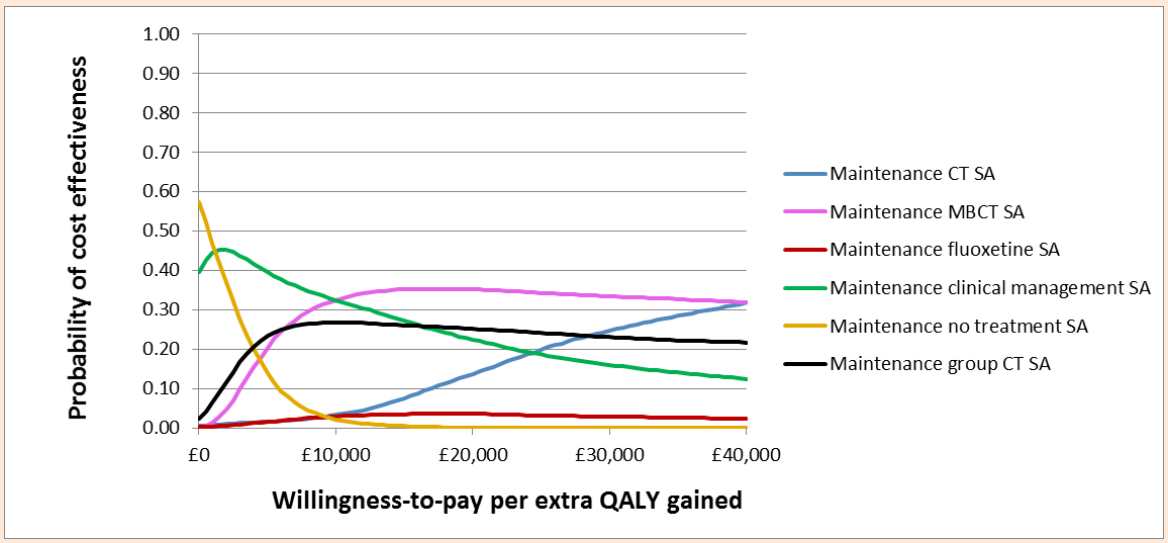
Update 2017

costs and QALYs versus clinical management per 1,000 adults. Base-case and sensitivity analysis



The probability of each option being cost-effective in the sensitivity analysis that included MBCT and group CT at various levels of WTP per QALY gained is shown in Figure 35.

Figure 35. Cost-effectiveness acceptability curves of interventions for people at high risk of relapse who remitted following acute psychological treatment and who experienced more severe depression if they relapsed – sensitivity analysis



When the number of individual CT sessions was reduced to 4, then individual CT became the most cost-effective treatment option, above MBCT and group CT, even if the preventive result of psychological interventions was assumed to last only one year.

13.3.54 People at high risk of relapse who remitted following acute combination treatment and who experienced more severe depression if they relapsed

The most cost-effective maintenance treatment option for people at high risk of relapse who remitted following acute combination treatment (represented by CBT and fluoxetine) was maintenance antidepressant treatment alone, with a high probability of being cost-effective that reached 0.95 at the NICE lower cost-effectiveness threshold of £20,000/QALY. This was

Update 2017

followed by maintenance combination therapy; the latter was more effective than maintenance antidepressant treatment with an ICER of £74,519/QALY. Psychological intervention plus clinical management (antidepressant drug tapering) was marginally less cost-effective than clinical management (antidepressant drug tapering) alone, as its ICER versus clinical management reached £20,995/QALY. The base-case results of the analysis are shown in Table 287. Results of deterministic analysis were similar.

Table 287: Results of economic modelling: interventions for people at high risk of relapse who remitted following acute combination treatment and who experienced more severe depression if they relapsed (mean values from probabilistic analysis)

Maintenance treatment option	Mean /person		ICER (£/QALY)	NMB (£) /person	Prob best[1]	Mean ranking
	QALY	Cost				
Combination therapy	6.732	5,937	74,519	128,694	0.04	2.33
AD alone (fluoxetine)	6.721	5,134	473	129,281	0.95	1.06
Psychological intervention (CBT) & clinical management (AD tapering)	6.709	5,877	Dominated	128,308	0.01	3.36
Clinical management (AD tapering)	6.673	5,112		128,344	0.00	3.25

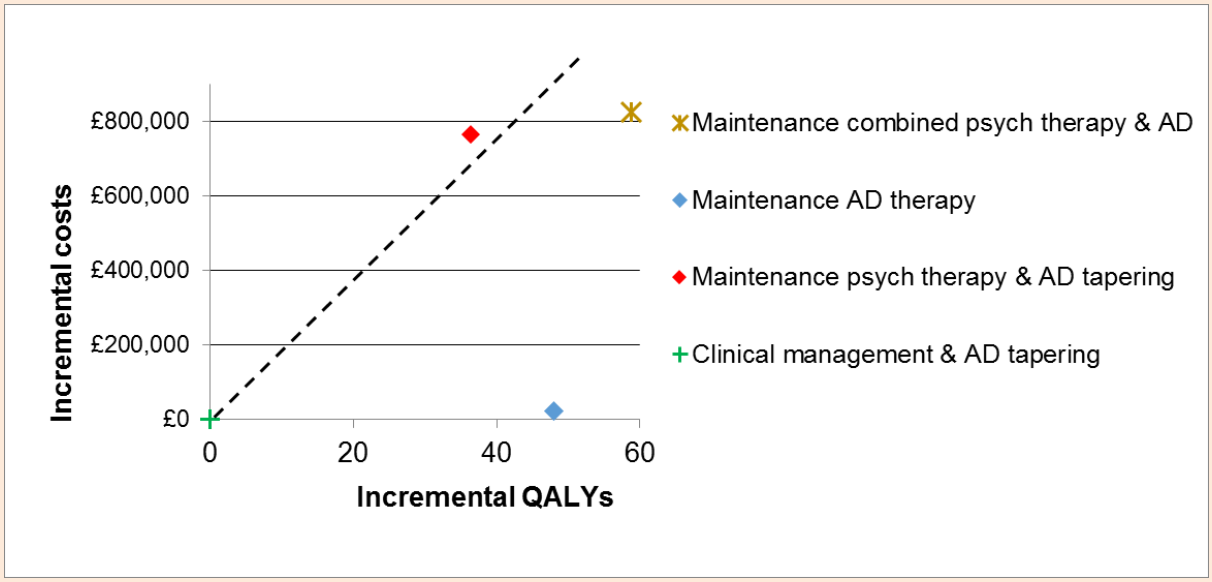
Notes:
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
AD: antidepressant; CBT: cognitive behavioural therapy; Prob: probability

Figure 36 provides the cost effectiveness plane of the analysis. Each intervention is placed on the plane according to its incremental costs and QALYs compared with clinical management and antidepressant drug tapering, which is placed at the origin. The slope of the dotted line indicates the NICE lower cost effectiveness threshold, suggesting that maintenance antidepressant treatment, alone or combined with psychological therapy, are cost-effective compared with clinical management and antidepressant drug tapering for people at high risk of relapse who remitted following acute combined treatment (since they both lie on the right side of the dotted line). Maintenance psychological therapy alone is less cost-effective than clinical management and antidepressant drug tapering, lying on the left of the dotted line of cost effectiveness.

Figure 36 Cost effectiveness plane of maintenance interventions for people at high risk of relapse who remitted following acute combined psychological and pharmacological treatment and who experienced more severe depression if

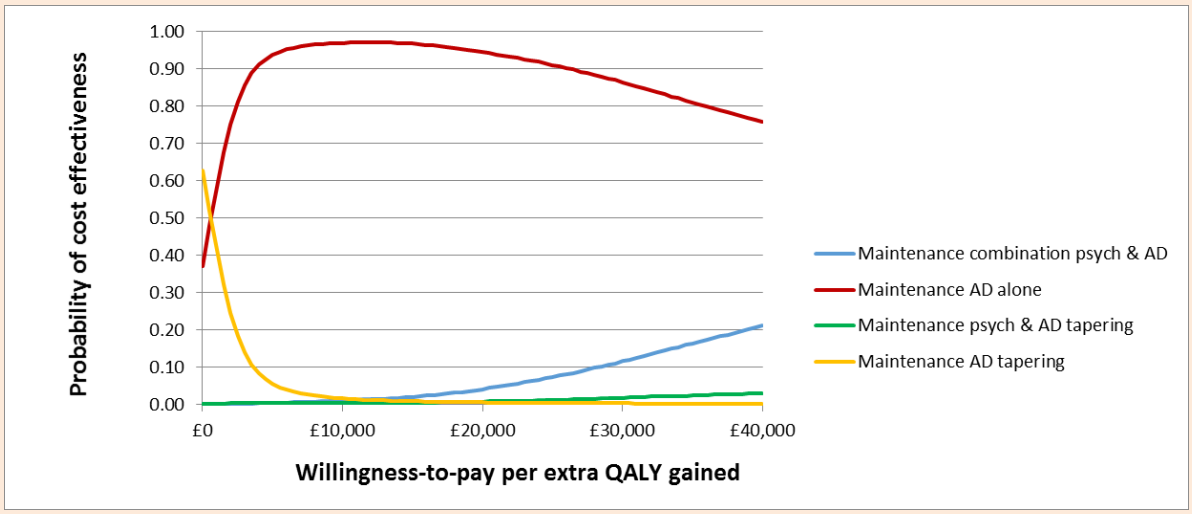
Update 2017

they relapsed – incremental costs and QALYs versus clinical management and antidepressant drug tapering per 1,000 adults.



The probability of each intervention being cost-effective in people at high risk of relapse to more severe depression who remitted following acute psychological treatment at various levels of WTP per QALY gained is shown in Figure 37.

Figure 37. Cost-effectiveness acceptability curves of interventions for people at high risk of relapse who remitted following acute combination treatment and who experienced more severe depression if they relapsed



In deterministic sensitivity analysis, if the number of previous episodes increased from 3 to 5, maintenance antidepressant treatment and maintenance combination treatment remained the best and second best option, respectively. Psychological therapy combined with clinical management (drug tapering) became more cost-effective than clinical management (drug tapering) alone.

When the future relapse episodes were assumed to be less severe or when alternative (higher) utility values were used for more severe depression, maintenance antidepressant treatment remained the most cost-effective option, followed by clinical management (drug tapering), and then combination maintenance treatment as third most cost-effective option; psychological therapy combined with clinical management (drug tapering) was the least cost-effective option.

Update 2017

Assuming a zero cost for clinical management or changing the cost of relapse by ± 50% had no impact on the results of the base-case analysis.

Assuming that psychological therapy comprised 4 sessions and retained its effect, maintenance antidepressant treatment and maintenance combination treatment remained the best and second best option, respectively. Psychological therapy combined with clinical management (drug tapering) became more cost-effective than clinical management (drug tapering) alone, even if its preventive effect was assumed to last only one year.

When the preventive effect of psychological treatment was assumed to last only one year (but psychological therapy comprised 10 sessions), maintenance antidepressant treatment remained the most cost-effective option, followed by clinical management (drug tapering), combination maintenance treatment, and, finally, psychological intervention combined with clinical management (antidepressant tapering).

A threshold analysis revealed that, for combination therapy to become the most cost-effective option, the number of sessions of CBT would need to fall at 4 and at the same time the number of previous depressive episodes would have to reach 7.

13.46 Discussion – conclusions, strengths and limitations of economic analysis

The guideline economic analysis assessed the cost effectiveness of a range of pharmacological and psychological interventions for the maintenance treatment of adults with depression that is in remission treated predominantly in primary care. The analysis considered appropriate interventions for adults with depression according to the acute treatment that led to remission of their most recent depressive episode, and also according to their risk for future relapses, as determined by their number of previous depressive episodes. Conclusions from the guideline economic analysis may be relevant to people in secondary care, especially given that clinical evidence was derived almost exclusively from studies conducted in secondary care settings (however, it needs to be noted that costs utilised in the guideline economic model were mostly relevant to primary care).

In people at medium risk of relapse who have remitted following acute pharmacological treatment (SSRIs, SNRIs, TCAs or mirtazapine) and who are expected to experience less severe depression if they relapse, maintenance pharmacological treatment is highly unlikely to be cost-effective compared with clinical management plus antidepressant drug tapering (probability of drugs being cost-effective ranging from 0.07 for SNRIs to 0.30 for SSRIs at the NICE lower cost-effectiveness threshold of £20,000/QALY). Maintenance pharmacological treatment, in particular with SSRIs, appears to be cost-effective if future episodes are more severe and as the risk of relapse increases (reflected in a higher number of previous episodes). This finding is explained by the low benefit-to-harm ratio of antidepressants in this population: the absolute risk of relapse is low (0.103 in the first year in people with one previous episode without maintenance drug treatment), the deterioration in HRQoL due to future relapse is milder (as relapses are less severe), and the risk of developing common side effects due to antidepressants and thus experiencing a utility decrement is relatively high (ranging from 0.117 with SSRIs to 0.163 with mirtazapine). However, as the number of previous episodes increases, the absolute risk of relapse increases and the preventive effect of maintenance drug treatment is enhanced; moreover, if relapses are more severe, the decrement in HRQoL resulting from each relapse increases, and the preventive effect of drugs has a larger (positive) impact on HRQoL. Consequently, the harms of maintenance drug treatment (side effects) are offset by its benefits (reduction in the number of relapses and larger improvement in HRQoL from prevention of relapses).

In people at high risk of relapse who have remitted following acute pharmacological treatment and who are expected to experience more severe depression if they relapse, the combination of MBCT with clinical management (antidepressant drug tapering) appears to be

Update 2017

the most cost-effective option with quite high certainty (probability of being cost-effective 0.83 at the NICE lower cost-effectiveness threshold of £20,000/QALY). MBCT combined with antidepressant treatment is the second most cost-effective treatment option, followed by maintenance antidepressant treatment. The combination of MBCT with clinical management (antidepressant drug tapering) remained cost-effective in sensitivity analysis that included maintenance group CT combined with antidepressant treatment. MBCT plus clinical management (antidepressant drug tapering) appeared to be the most cost-effective option under a range of scenarios explored in sensitivity analysis. However, if the preventive effect of MBCT lasts only one year, then the combination of MBCT plus antidepressant treatment becomes the most cost-effective intervention, followed by MBCT plus clinical management (antidepressant tapering), then antidepressant treatment alone, and, finally, clinical management and antidepressant drug tapering. Results are driven by the effectiveness of MBCT combined with the low intervention cost of (group-delivered) MBCT.

In people at medium risk of relapse who have remitted following acute psychological treatment and who are expected to experience less severe depression if they relapse, clinical management appears to be the most cost-effective intervention (with a probability of 0.58 at the NICE lower cost-effectiveness threshold of £20,000/QALY), followed by no treatment. Maintenance psychological treatment (CT) consisting of 10 individual hourly sessions appears to be the third most cost-effective option among those assessed in this analysis. However, if the preventive effect of CT can be achieved with 4 individual hourly sessions so that the intervention cost is greatly reduced, then CT appears to become the most cost-effective maintenance treatment option among those assessed in this population, provided that its relapse preventive effect lasts two years. The results are driven by the uncertainty characterising the clinical efficacy model input parameters, the relatively high cost of individual CT and the relatively low risk of relapse characterising the study population.

In people at high risk of relapse who have remitted following acute psychological treatment and who are expected to experience more severe depression if they relapse, clinical management appears to be the most cost-effective option (with a probability of 0.39 at the NICE lower cost-effectiveness threshold of £20,000/QALY) followed by maintenance CT. In sensitivity analysis that included group CT and MBCT, MBCT became the most cost-effective option, while group CT was the fourth most cost-effective option behind clinical management and maintenance CT. If the preventive effect of individual CT can be achieved with 4 hourly sessions, then CT becomes the most cost-effective option among all interventions assessed (including MBCT and group CT), even if its relapse preventive effect lasts only one year. The results are driven by the uncertainty characterising the clinical efficacy model input parameters and the relatively high cost of individual CT.

In people at high risk of relapse who have remitted following combined pharmacological and psychological acute treatment and who are expected to experience more severe depression if they relapse, maintenance pharmacological treatment alone appears to be the most cost-effective intervention followed by combination therapy. The probability of pharmacological treatment alone being the most cost-effective maintenance treatment option in this population is very high (0.95 at the NICE lower cost-effectiveness threshold of £20,000/QALY). It is noted that combination therapy is the most effective intervention; however, it has also a high intervention cost, mainly driven by the cost of maintenance psychological therapy, which comprises 10 individual CBT sessions. Nevertheless, even if the preventive effect of combined pharmacological and psychological therapy can be achieved with 4 individually delivered hourly sessions of CBT, meaning that the cost of combination therapy is greatly reduced, maintenance pharmacological treatment remains the most cost-effective treatment option. According to threshold analysis, combination therapy becomes the most cost-effective option when the psychological treatment component consists of 4 individual hourly sessions, and the population has at least 7 previous depressive episodes, so that the risk of relapse is increased and the impact of the preventive effect of combination therapy is enhanced. Psychological therapy plus clinical management (antidepressant drug tapering) appears to be marginally less cost-effective than clinical

Update 2017

management (drug tapering) alone; its relative cost effectiveness versus clinical management improves when the number of previous episodes (and therefore the risk of future relapses) increases and when psychological therapy comprises 4 individual sessions (rather than 10). Results are driven by the high effectiveness of antidepressant therapy alone or in combination with psychological therapy and the high cost of psychological therapy if it consists of 10 individual CBT sessions.

Results of the economic analysis were overall robust to different scenarios explored through sensitivity analysis. In general, the relative cost effectiveness of more effective interventions improved when the risk of relapse (as reflected in number of previous episodes) increased, because their preventive effect had a greater impact (as a higher number of future relapses was avoided), and associated cost-savings offset the maintenance intervention costs. The cost effectiveness of individual psychological interventions improved when the number of sessions was reduced, provided that their relapse preventive effect was fully retained.

The economic analysis enabled estimation of the cost effectiveness of appropriate interventions for adults at medium risk of relapse (1-2 previous depressive episodes) to less severe depression and those at high risk of relapse (3+ previous depressive episodes) to more severe depression and allowed exploration of changes in the relative cost effectiveness of interventions with increasing number of previous depressive episodes, thus with increasing risk of relapse. The analysis also allowed consideration of cost effectiveness of interventions depending on the type of acute treatment (i.e. pharmacological, psychological or combined) people had received that led to remission of their most recent depressive episode.

Most available efficacy data were not specific to the risk of relapse of the study population, as determined by the number of previous depressive episodes. However, most studies reported some indicator of the number of previous episodes experienced by the study participants, such as mean or median number of previous episodes or the minimum number of previous episodes required as an inclusion criterion. This allowed categorisation of the study participants in each study as being at low, moderate or high risk of relapse. Some interventions considered in the guideline systematic review were tested exclusively on high risk populations, so the respective evidence was utilised only in populations at high risk of relapse in the economic analysis. Also, available evidence did not focus on the severity of depression; therefore distinguishing future episodes of depression into less and more severe in the economic model was exclusively determined by the utility value attached to future depressive episodes (all of which, in each cohort examined, had to be either less severe or more severe).

The analysis utilised clinical effectiveness parameters derived from NMAs conducted separately for each population of interest. This methodology enabled evidence synthesis from both direct and indirect comparisons between interventions, and allowed simultaneous inference on all treatments examined in pair-wise trial comparisons while respecting randomisation (Caldwell et al., 2005; Lu & Ades, 2004). However, due to lack of relevant data from primary care settings, efficacy data were derived from RCTs conducted in secondary care and thus may not be directly relevant to the study population. Furthermore, the quality and limitations of RCTs considered in the NMAs have unavoidably impacted on the quality of the economic model clinical input parameters. For example, economic results may be have been affected by reporting and publication bias.

A number of RCTs included in the guideline systematic review compared psychological interventions versus TAU, and were thus not possible to include in the main networks constructed for each population. Nevertheless, after identifying what constituted TAU in each cohort, these studies were possible to include in NMA and economic sensitivity analyses and to consider as additional treatment options for relevant populations.

The NMAs estimated hazard ratios for each intervention versus the baseline comparator (pill placebo), which was the most appropriate output given the underlying Weibull distribution

Update 2017

characterising the risk of relapse. These hazard ratios were subsequently applied onto the baseline risk of relapse over the first 2 years of the analysis, in order to calculate the specific risk of relapse associated with each intervention and each population assessed in the economic analysis.

The relapse preventive effect of all interventions assessed in the model (pharmacological, psychological and combined) was assumed to last over 2 years from initiation of maintenance treatment in the base-case analysis. However, evidence on the longer-term effects of maintenance psychological interventions is limited and suggests that the effect of psychological interventions may actually diminish over time. Nevertheless, a scenario under which the effect of psychological interventions lasted only over the first year form initiation of maintenance therapy was tested in sensitivity analysis.

The baseline risk of relapse and the probability of recovery over time were estimated based on a review of naturalistic studies. Available data suggested that both parameters were characterised by a Weibull distribution, in which the events rates are proportional to a power of time. The economic analysis incorporated Weibull distribution characteristics for both input parameters, derived from available evidence, thus enabling a better representation of the course of depression over time. The increase in the risk of future relapses imposed by each additional depressive episode experienced by people with depression was also factored in the economic analysis by the means of a hazard ratio of relapse with every additional depressive episode.

The time horizon of the analysis was 10 years, which was considered by the GC long enough to capture longer-term benefits and costs (including cost-savings) associated with the preventive effect of interventions assessed.

Utility data used in the economic model were derived from a systematic review of studies reporting utility data for depression-related health states that were generated using the EQ-5D and the UK population tariff, as recommended by NICE.

NHS and PSS costs incurred by adults with depression that is in remission or in a depressive episode were derived from a large (N=88,935) naturalistic study that aimed to estimate health service use and costs associated with non-remission in people with depression using data from a large primary care UK general practice research database (Byford et al., 2011). The study utilised data collected between 2001 and 2006 and, although not recent, was considered the best source of cost information for the study population as it provided detailed data of healthcare resource use relating to the primary care treatment of adults with depression in the UK. Resource estimates and unit costs were updated with 2016 cost data and supplemented with further evidence according to GC expert advice, where appropriate, to reflect current routine practice in the UK NHS.

Maintenance treatment discontinuation has not been explicitly considered in the model structure. However, the clinical efficacy data utilised in the analysis have implicitly accounted for discontinuation, as an intension-to-treat approach was adopted in the guideline data extraction and meta-analysis. Moreover, the probabilistic model did assume that a percentage of people in the cohort might have not completed treatment or they might have had less than perfect compliance, so a less than full intervention cost has been assumed for these people.

The impact of common side effects from maintenance antidepressant treatment alone or in combination on HRQoL and costs associated with their management was incorporated in the economic analysis. No side effects were considered for people receiving non-pharmacological interventions; however, people receiving non-pharmacological treatments for depression are also expected to experience a range of events such as headaches, nausea or vomiting, etc. Therefore, the economic analysis may have overestimated the impact of common side effects from antidepressants relative to other treatments and thus underestimated their relative cost effectiveness. On the other hand, other less common side

Update 2017

effects associated with treatment with antidepressants (such as upper gastrointestinal bleeds and falls) were not considered in the economic model. Such side effects result in considerable reduction in HRQoL and high costs for their management; nevertheless, they are relatively rare and therefore their omission is unlikely to have significantly impacted on the model results, although it is acknowledged as a limitation that has potentially overestimated the cost effectiveness of drugs or combined interventions with a pharmacological intervention element relative to other maintenance treatments.

13.58 Overall conclusions from the guideline economic analysis

In people at medium risk of relapse who have remitted following acute pharmacological treatment and who are expected to experience less severe depression if they relapse, maintenance pharmacological treatment with the same drug they had received as acute treatment over 2 years is not cost-effective versus clinical management (antidepressant tapering) due to the high harm-to-benefit ratio of maintenance drug treatment in this population. The cost effectiveness of maintenance drug treatment increases as the severity of depression increases and as the risk for future relapses, as determined by the number of previous episodes, increases.

In people at high risk of relapse who have remitted following acute pharmacological treatment and who are expected to experience more severe depression if they relapse, maintenance treatment with MBCT in combination with clinical management (antidepressant drug tapering) appears to be the most cost-effective option with high probability, followed by combination of MBCT with antidepressant treatment. Maintenance antidepressant treatment alone is more cost-effective than clinical management with antidepressant tapering. However, if the preventive effect of MBCT lasts only one year, then the combination of MBCT plus antidepressant treatment becomes the most cost-effective intervention, followed by MBCT plus clinical management (antidepressant tapering), antidepressant treatment alone, and, finally, clinical management and antidepressant drug tapering.

In people at medium risk of relapse who have remitted following acute psychological treatment and who are expected to experience less severe depression if they relapse, maintenance high intensity CT (comprising 10 individual hourly sessions) does not appear to be cost-effective, and clinical management or no treatment should be preferred instead. However, if the preventive effect of CT can be achieved with 4 individual hourly sessions so that the intervention cost is greatly reduced, then maintenance CT becomes cost-effective provided that its relapse preventive effect lasts two years.

In people at high risk of relapse who have remitted following acute psychological treatment and who are expected to experience more severe depression if they relapse, maintenance CT comprising 10 individual hourly sessions and with an effect that lasts two years is marginally less cost-effective than clinical management. Maintenance CT consisting of 4 individual hourly sessions (provided that it can achieve the same effect as CT comprising 10 individual sessions over a minimum of one year) is more cost-effective than clinical management. MBCT also appears to be a cost-effective option for this population, although less cost-effective than 4 individual hourly sessions of CT (provided that its effect is equal to that of CT comprising 10 individual sessions).

In people at high risk of relapse who have remitted following combined pharmacological and individual psychological acute treatment and who are expected to experience more severe depression, maintenance pharmacological treatment alone is highly likely the most cost-effective treatment option. Combination therapy is the most cost-effective option if it includes a less intensive psychological component (e.g. 4 individual hourly sessions that retain the effect of 10 sessions), and the population's risk of relapse is quite high, as determined by a higher number (at least 7) of previous depressive episodes. Maintenance individual psychological therapy plus clinical management (drug tapering) becomes potentially more cost-effective than clinical management alone as the number of previous episodes (and thus

Update 2017

the risk of relapse characterising the study population) increases or if the number of individual sessions is reduced to 4 (provided that the effect of 10 individual sessions can be achieved for a minimum of one year).

Overall, the relative cost effectiveness of more effective interventions improves when the risk of relapse (as reflected in number of previous episodes) increases, because their preventive effect has a greater impact (as a higher number of future relapses is avoided), and associated cost-savings offset the maintenance intervention costs.

Conclusions from the guideline economic analysis refer mainly to people with depression who are predominantly treated in primary care; however, they may be relevant to people in secondary care as well, especially given that clinical evidence was derived almost exclusively from studies conducted in secondary care settings (however, it needs to be noted that costs utilised in the guideline economic model were mostly relevant to primary care).

Update 2017

14 Economic modelling: cost effectiveness of interventions for the treatment of new depressive episodes in adults

14.1 Introduction – objective of economic modelling

The choice of initial treatment for adults with a new depressive episode was identified by the GC and the guideline health economist as an area with potentially major resource implications. Although existing economic evidence in this area is quite extensive, no study has currently assessed the relative cost effectiveness of the whole range of available interventions for people with a new episode of depression in the UK. The guideline network meta-analysis (NMA) synthesised available clinical evidence in order to inform an economic model, developed to assess the relative cost effectiveness between all effective interventions considered in the NMA. Based on the above considerations, an economic model was developed to assess the relative cost effectiveness of pharmacological, psychological, physical and combined interventions for adults with a new episode of depression in the UK.

The purpose of the model is to assess the best approach for treatment of a new episode of depression up to its (potential) resolution and includes a two-year follow-up, in order to incorporate cost-effective maintenance therapy aiming at preventing relapse in people who have remitted following acute treatment. However, people with depression may experience multiple recurrent episodes, which have not been incorporated in the acute treatment model structure. The consequences (costs and impact on health-related quality of life [HRQoL]) of recurrent depressive episodes in the longer-term have been considered in a separate model that was developed to assess the cost effectiveness of interventions for depression aiming at preventing relapse in adults with depression that is in remission. The economic analysis of interventions for relapse prevention is described in Chapter 13.

Update 2017

14.2 Methods

14.2.1 Population

The study population of the economic model comprised adults with depression initiating treatment for a new episode in primary care. This was decided because the majority of adults with a new episode of depression are treated in primary care in routine UK practice. Two populations were considered: adults with a new episode of less severe depression and adults with a new episode of more severe depression. The definition of less severe and more severe depression was the same as that used to classify RCTs in the two respective NMAs undertaken to estimate the acceptability and effectiveness of interventions for the treatment of a new episode of depression, which informed the economic analysis. The definition of less severe and more severe depression is provided in Chapter 7, section 7.2. The study population had no physical comorbidities, psychotic symptoms, complex or chronic depression in accordance with the inclusion criteria of the systematic review of RCTs that informed the NMAs.

People in the economic analysis were assumed to be experiencing their first depressive episode if they had less severe depression and their fourth depressive episode if they had more severe depression, to cover a range of adults with a new episode of depression presenting in routine clinical practice. The number of previous episodes determined the study population's risk of relapse following remission of the current episode.

The age of the cohorts considered in the economic model was determined by the mean age of onset of depression in adults and the number of the current new episode for which treatment was received.

Kessler, Berglund et al. (2005) reported the results of a national comorbidity household survey in the US, according to which the median age-of-onset of depression was 32 years (interquartile range 19-44 years). In a Swedish longitudinal cohort study of 3,563 people followed up for 30-49 years, the median age at first onset of depression was reported to be around 35 years (Mattisson, Bogren et al. 2007). A large (n=20,198) Scottish family-based population study designed to identify the genetic determinants of common diseases, including major depression disorder, reported a mean age of onset of major depressive disorder of 31.7 years (SD 12.3 years) among 2,726 participants that met DSM-IV criteria for current and/or past major depression disorder (Fernandez-Pujals, Adams et al. 2015). On the other hand, Andrade, Caraveo-Anduaga et al. (2003) did a review of results of community epidemiological surveys on major depressive episodes that were carried out in 10 countries in America, Europe and Asia (the UK was not included in these countries); the authors reported a median age of onset of major depression in the early to mid-twenties in all countries other than Japan (late twenties) and the Czech Republic (early thirties). Based on this evidence and following GC expert advice, the age of onset of major depression in the study population was set at 32 years.

According to the GC expert opinion, the mean interval between 2 consecutive depressive episodes in people who experience relapses is about 2 years. Therefore, for modelling purposes, people with a new episode of less severe depression were assumed to be 32 years of age (as this was their first episode) and people with more severe depression were assumed to be 38 years of age (as this was their fourth episode).

The percentage of women in each cohort were estimated to be 56%, based on weighted epidemiological data on depressive episodes reported in the most recent adult psychiatric morbidity household survey conducted in England (McManus, Bebbington et al. 2016).

Determining the age and gender mix of the cohorts was necessary in order to estimate mortality risks in the model.

Update 2017

14.2.2 Interventions assessed

The range of interventions assessed in the economic analysis was determined by the availability of relevant clinical data synthesised in the NMA. Selection of interventions followed a step-wise approach, in which classes of interventions were selected for consideration first, followed by selection of interventions within each selected class. The selection of interventions was made based on the following algorithm and criteria:

- The economic analysis on each population (i.e. people with less severe depression and people with more severe depression) assessed only interventions that were included in the respective (in terms of study population) NMAs.
- For each population, only classes of interventions with available NMA data on the outcomes of discontinuation and response in completers were included in the analysis, as these outcomes were essential in order to populate the economic model. The full list of NMA outcomes considered in the economic analysis are reported in section 14.2.5.
- Tricyclic antidepressants (TCAs) were excluded from the economic analyses because the GC did not consider them as suitable first line treatments for a new episode of depression in primary care (which was the setting adopted in the economic analyses) due to their side effect profile.

Once the classes of interventions for inclusion in the economic analysis were determined, individual interventions were selected as representatives within each class, where data on

more than one intervention were available. The final selection of interventions within each class was based on judgement, using a number of criteria:

- relative effectiveness: more effective interventions within the class were better candidates for selection; where more than one interventions were selected within one class, to reflect different resource use intensity, selection of interventions with a range of effects (as indicated by ranking) was attempted
- availability of discontinuation and completers' response NMA data for specific interventions
- number of people tested on each intervention
- availability of interventions within the NHS: more commonly used interventions had a priority over less commonly used interventions
- resource use intensity: this was not a criterion if interventions had a similar resource intensity and thus similar intervention cost; however, where the format of interventions within a class differed, resulting in considerably different intervention costs (for example, individually versus group delivered interventions), then interventions with different modes of delivery within a class were selected for inclusion in the economic analysis.

Assessment of the cost effectiveness of interventions and classes that were not possible to include in the economic analysis due to lack of data on relevant outcomes was based on comparison of their relative effects and intervention resource use with interventions and classes that were included in the economic analysis.

In addition to active interventions, the economic model also considered non-specific clinical management by GPs, as a benchmark treatment option, which, in terms of effectiveness, was reflected in RCT pill placebo arms. Clinical management was considered as an option for both study populations. Based on the above criteria, the following interventions were included in the economic analysis for each study population:

Adults with less severe depression:

- pharmacological interventions: citalopram; mirtazapine
- psychological interventions: behavioural activation (BA); Coping with Depression course (group); cognitive behavioural therapy (CBT) individual (over 15 sessions); CBT group (under 15 sessions); interpersonal psychotherapy (IPT); short term psychodynamic psychotherapy (PDPT) individual; non-directive counselling; computerised CBT with support; computerised CBT without support; psychoeducational group programme
- physical interventions: physical exercise programme
- combined interventions: CBT individual (over 15 sessions) + citalopram; IPT + citalopram; short term PDPT individual + citalopram; physical exercise programme + sertraline
- clinical management, reflected in pill placebo RCT arms

Adults with more severe depression:

- pharmacological interventions: sertraline; mirtazapine
- psychological interventions: BA; CBT individual (over 15 sessions); CBT group (under 15 sessions); short term PDPT individual; non-directive counselling; computerised CBT without support
- physical interventions: physical exercise programme
- combined interventions: CBT individual (over 15 sessions) + sertraline
- clinical management, reflected in pill placebo RCT arms

Update 2017

14.2.31 Model structure

A hybrid decision-analytic model consisting of a decision-tree followed by a three-state Markov model was constructed using Microsoft Office Excel 2013. The model estimated the total costs and benefits associated with provision of effective treatment options in two cohorts of adults with a new episode of less severe and more severe depression, respectively. The structure of the model, which aimed to simulate the course of depression and relevant clinical practice in the UK, was also driven by the availability of clinical data.

According to the model structure, hypothetical cohorts of adults with a new episode of depression were initiated on each of the treatment options assessed, as appropriate, according to their level of symptom severity. People in each cohort either completed treatment or discontinued early due to intolerable side effects or other reasons. The duration of a full course of initial treatment was 12 weeks for drugs and clinical management; the duration of psychological and physical interventions varied by intervention (ranging between 6 and 16 weeks). The duration of combined interventions was determined by the component with the longest duration. For practical purposes of estimation of QALYs it was assumed that all interventions lasted 12 weeks, without this assumption affecting resource use associated with each intervention. People who discontinued an active treatment early were assumed to switch to a mixture of available treatments for depression or no treatment; people who discontinued clinical management were assumed to move to no treatment. The mixture of available treatments following discontinuation was assumed to have the effectiveness of the baseline treatment (clinical management) and a mean management cost of people in a depressive episode. No treatment was assumed to have the effectiveness of wait list and zero cost. The proportion of people moving to no treatment after discontinuation of the active treatment equalled the probability of discontinuation (and moving to no treatment) under clinical management.

Following completion of initial treatment or early discontinuation and switch to a mixture of treatments or no treatment, people in each branch of the model either remitted, or responded to treatment without reaching remission, or failed to meet criteria for response. These 3 states (response reaching remission; response not reaching remission; no/inadequate response) were the endpoints of the decision-tree component of the model. From that point on, all people entered the Markov component of the model, which consisted of 3 states: remission (no depressive episode); depressive episode (either due to persistence of the current episode or due to relapse); and death. People who were in remission at the decision-tree endpoint moved to the remission state; those who did not meet criteria for response at the decision-tree endpoint moved to the depressive episode state; and those who responded but did not meet criteria for remission were assumed to either remit (thus moving to the remission state of the Markov model) or remain in a depressive episode (thus moving to the depressive episode state of the Markov model).

The Markov model was run in yearly cycles with a half-cycle correction being applied. In each model cycle, people entering the Markov component of the model could either remain in the same ‘entrance’ state, move between the remission and the depressive episode states, or move to the death state (absorbing state). People with more severe depression who remitted from their 4th episode following treatment completion were assumed to receive optimal relapse prevention treatment, as appropriate, depending on the acute treatment that led to remission, as determined by the guideline recommendations on relapse prevention treatments included in Chapter 11. Details on the specific maintenance treatment received by each cohort are provided at the end of this section. Maintenance antidepressant treatment lasted 2 years; maintenance psychological treatment lasted 1 year. Benefits of all treatments were assumed to be enjoyed over 2 years, according to available evidence on pharmacological and psychological interventions aiming at relapse prevention and the GC expert opinion. People with less severe depression who remitted from their 1st episode following treatment completion were assumed to receive no relapse preventive treatment,

Update 2017

apart from 3 extra GP visits in the first year and 1 extra GP visit in the second year they spent in the Markov remission state.

The duration of the Markov model component was 2 years, to enable the full costs and effects of a course of treatment for depression (including acute and, if appropriate, maintenance treatment) to be modelled. Thus, the total time horizon of the economic analysis was 12 weeks of acute treatment (decision-tree) plus 2 years of follow up which included maintenance treatment, as appropriate, for people who remitted following successful acute treatment (Markov model).

The baseline risk of relapse in the Markov remission state depended on the time (one or two years) people spent in this state (the longer people stayed in remission, the lower their risk of relapse) and their number of previous episodes (the higher the number of their previous episodes, the higher their risk of relapse). Therefore, over the 2 years of the Markov component of the model, the risk of relapse experienced by each cohort was determined by their baseline risk of relapse and the efficacy of the (potential) maintenance treatment option received by each cohort. If people relapsed during this period of 2 years, maintenance treatment was discontinued and the preventative benefit of maintenance treatment ceased at the point of relapse.

The probability of remission for each cohort in the depressive episode state depended on the time (one or two years) people spent in this state (the longer people stayed in the depressive episode, the lower their probability of remission) and the severity of depression (less or more severe).

Within the remission and depressive episode states, people entered tunnel states, so that the time they remained in every state (one or two years) could be estimated and a time-dependent probability of relapse or remission, respectively, could be applied.

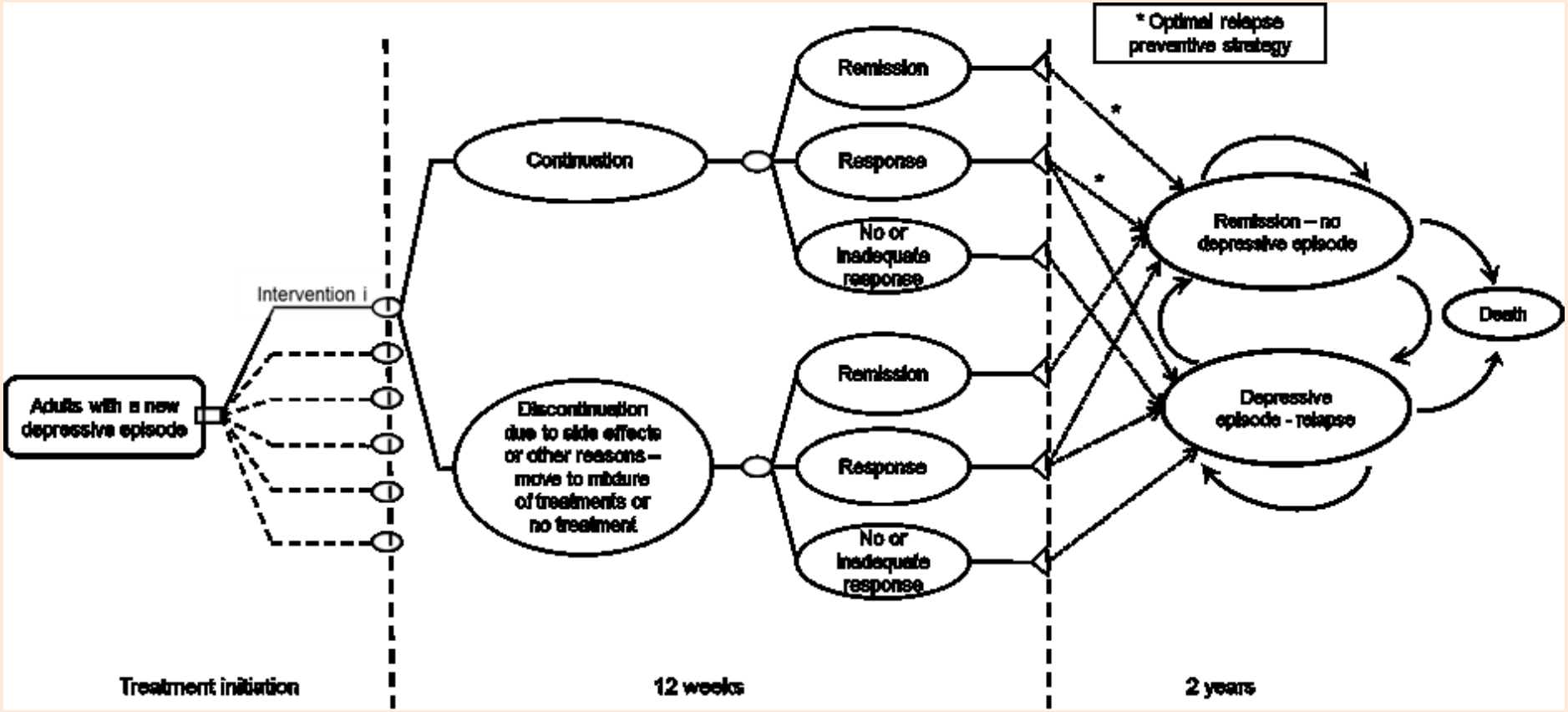
Death was not considered in the acute part of the model. Although the mortality risk in people with depression is higher than that of people in the general population (Cuijpers, Vogelzangs et al. 2014), suicide (which is the main cause of death in adults with a new episode of depression) is a rare outcome in trials, and there are no substantial differential data on suicide between treatments. The GC expressed the view that consideration of suicide in the acute part of the model would have no significant impact on the relative cost effectiveness between different treatments, and therefore death was considered only in the Markov component of the economic model, for which more relevant, long-term data were available.

Side effects from medication were considered in the model in 2 ways: people who discontinued pharmacological treatment due to side effects were assumed to experience a reduction in their HRQoL over 5 weeks (approximately over the period they were receiving antidepressant treatment) and to incur one extra GP visit. People who completed antidepressant treatment were assumed to experience common antidepressant side effects (such as headaches, nausea, agitation, sedation, sexual dysfunction) resulting in a reduction in their HRQoL over the period they received antidepressant treatment (i.e. 12 weeks of acute antidepressant treatment plus 2 years for those receiving maintenance antidepressant treatment). Moreover, they were assumed to incur extra costs for the management of their side effects, which comprised GP visits and pharmacological treatment.

The structure of the economic model for interventions for people with a new episode of depression is shown in Figure 38.

Update 2017

Figure 38. Schematic diagram of the structure of the economic model of treatment of new depressive episodes in adults



Update 2017

Relapse-preventive interventions received by people with depression who remitted following successful acute treatment

People with more severe depression in their 4th episode who remitted following successful acute treatment moved on to appropriate relapse preventive intervention. Table 288 shows the type of maintenance people received according to the acute treatment that led to remission of the depressive episode.

Table 288: Type of maintenance therapy received by people in the model according to the acute treatment that led to remission of the depressive episode

Acute treatment	Subsequent maintenance treatment aiming at relapse prevention
More severe depression (remission of fourth depressive episode)	
Sertraline	80%: 2 years of maintenance sertraline treatment 20%: group MBCT + drug tapering
Mirtazapine	80%: 2 years of maintenance mirtazapine treatment 20%: group MBCT + drug tapering
Behavioural activation	50%: 4 sessions of behavioural activation 50%: MBCT
CBT individual (over 15 sessions)	50%: 4 sessions of CBT individual 50%: MBCT
CBT group (under 15 sessions)	50%: group CBT 50%: MBCT
Short term PDPT individual	50%: 4 sessions of short term PDPT individual 50%: MBCT
Counselling	50%: 4 sessions of counselling 50%: MBCT
cCBT without support	50%: group CBT 50%: MBCT
Physical exercise programme	50%: group CBT 50%: MBCT
CBT individual (over 15 sessions) + sertraline	80%: 2 years of maintenance sertraline treatment 20%: 4 sessions of CBT individual + drug tapering
Clinical management	All: clinical management follow-up

Update 2017

14.2.49 Costs and outcomes considered in the analysis

The economic analysis adopted the perspective of the NHS and personal social services, as recommended by NICE (NICE 2014). Costs consisted of intervention costs (drug acquisition, staff time for provision of pharmacological, psychological, physical and combined therapies), including optimal maintenance treatments for relapse prevention in people who remitted, as well as costs associated with the further management of people who discontinued the initiated treatment, those who did not remit or people who relapsed following remission, which included drug acquisition, primary care, hospitalisation, outpatient visits, psychological therapies, and also accident and emergency visits. Costs of management of common side effects from antidepressants in people receiving pharmacological treatment and healthcare costs incurred by people in remission (potentially unrelated to the treatment of depression) were also considered in the analysis. The cost year was 2016.

The measure of outcome was the Quality Adjusted Life Year (QALY), which incorporated utilities associated with the health states of remission, response without reaching remission,

no or inadequate response, as well as utility decrements due to intolerable side effects and common (tolerable) side effects associated with antidepressant and combined treatment (both acute and maintenance).

14.2.54 Acceptability and efficacy data and methods of evidence synthesis

Acceptability and efficacy data for interventions considered in the economic modelling for a new episode of depression in people with less severe depression and people with more severe depression were derived from the NMAs of interventions for people with a new depressive episode that were undertaken for this guideline. Details on the methods and results of the NMAs, which were conducted in WinBUGS 1.4.3 (Lunn, Thomas et al. 2000, Spiegelhalter, Thomas et al. 2003) are provided in Chapter 17. In summary, binomial likelihood and logit models were used (Dias, Welton et al. 2011 [last updated 2013]), to allow estimation of odds ratios of each treatment versus baseline (which were different depending on outcome) for each outcome of interest, which were then applied onto the respective baseline risk of each outcome. For the economic analysis the first 100,000 iterations undertaken in WinBUGS were discarded and another 300,000 were run, thinned by 30, so as to obtain 10,000 iterations that populated the economic model.

Although, as discussed in Chapter 13, section 13.2.7, the probability of recovery in people with depression is reduced over time following a Weibull distribution, the logit model was considered appropriate to use for the estimation of relative effects between acute treatments expressed as odds ratios over a relatively short period of time.

For each population, the following parameters were obtained from the NMAs, expressed as odds ratios versus a selected baseline:

- discontinuation (for any reason)
- discontinuation due to side effects, in those discontinuing treatment
- response (reaching or not reaching remission) in those completing treatment
- remission in those completing treatment.

Update 2017

These data were combined with respective baseline risks for each outcome in people with less severe depression and in people with more severe depression, in order to estimate the probabilities of events of each intervention in each endpoint of the decision-tree component of the model, for each population of interest.

It needs to be noted that originally, the outcome of interest in order to populate the economic model with numbers of people remitting was remission conditional on response (i.e. probability of remission in those responding to treatment). However, the network constructed for this outcome in people with more severe depression was disconnected, and therefore relative effects between interventions of interest for this outcome were not possible to estimate for all comparisons. Moreover, the network constructed for this outcome in people with less severe depression was sparse and covered a limited number of interventions. For this reason, remission in those completing treatment was selected as an outcome instead, to allow, in combination with data on response in those completing treatment, calculation of numbers of people who responded and remitted. When running the probabilistic analysis, the number of people reaching remission was not allowed to exceed the number of people responding to treatment. In iterations where the probability of remission exceeded the probability of response, the number of people in remission was forced to equal that of people in response (so that all people who responded also remitted in those iterations). This approach was adopted in both economic analyses, for people with less severe depression and those with more severe depression.

Relative effects were obtained from the individual interventions' analyses, with the exception of discontinuation due to side effects in those discontinuing treatment. In these analyses, both in people with less and people with more severe depression, the class model effects

were utilised in the economic analysis, as individual intervention data were sparse and did not cover all interventions of interest; moreover, individual intervention data were based on small numbers randomised, compared with other analyses.

For some interventions considered in the economic analysis, relative effects were not available on one or more outcomes. In such cases, the intervention 'borrowed' the relative effect of another intervention of a similar type and format and with anticipated similar effect, ideally belonging to the same class, with the exception of remission in completers. When remission in completers' data were not available for a specific intervention, the probability of remission in responders, as estimated in the economic model, was used instead, if available for an intervention of a similar type and format and with anticipated similar effect, so that the impact of borrowing the effect from another intervention was more limited.

As discussed in section 14.2.6, for two of the outcomes (response in those completing treatment and remission in those completing treatment) the chosen baseline was pill placebo (reflected in clinical management). For the other two outcomes (discontinuation and discontinuation due to side effects in those discontinuing treatment) the selected baseline was the SSRI treatment included in the analysis (citalopram for less severe depression and sertraline for more severe depression).

The results of the network meta-analysis that were used to populate the economic model are provided in Table 289 for people with less severe depression and Table 290 for people with more severe depression.

It is noted that relative effects of treatments in the outcomes considered in the economic analysis may differ from those observed for the outcome of the SMD in those randomised in the clinical analysis. Possible explanations for this discrepancy include:

- Different studies may have been included in different analyses (depending on availability of reported outcome data in each study)
- There is a different way of accounting of drop-outs in each study outcome and each analysis: response in completers expresses improvement after excluding those who have discontinued treatment; the continuous scale data used in the SMD analysis, in most trials, express improvement in all people included in the trial, as some method of imputation has been used to estimate the effect in people who discontinued treatment. Last observation carried forward (LOCF) and multiple imputation account for people who discontinued treatment in a different way from baseline observation carried forward (BOCF). The dichotomous outcome of response in those randomised considers people who discontinued as non-responders. The NMA of response in those randomised includes a mixture of dichotomous response data (where people who discontinued were considered as non-responders) as a priority, in studies where such dichotomous data are available, and continuous data, where RCTs did not report dichotomous response data. Hence, the amount of and method of imputation for continuous data included in response in those randomised analyses have unavoidably affected the results of these analyses.

Update 2017

Table 289. Results of the NMAs that informed the economic analysis of interventions for a new depressive episode in people with less severe depression: odds ratios versus baseline for each outcome of interest

Intervention	Mean odds ratios versus baseline (95% credible intervals)			
	Discontinuation versus citalopram	Discontinuation due to side effects in those discontinuing versus SSRIs (respective class effects used)	Response in treatment completers versus pill placebo	Remission in treatment completers versus pill placebo
Citalopram	Baseline	Baseline	2.19 (1.29 to 3.73)	1.59 (0.56 to 4.54)
Mirtazapine	0.54 (0.12 to 2.50)	1.33 (0.10 to 19.99)	3.26 (0.67 to 16.53)	(probability of remission in responders borrowed from citalopram)
Behavioural activation	0.96 (0.42 to 2.23)	Not relevant	3.76 (1.68 to 8.40)	3.63 (1.52 to 8.88)
Coping with Depression course (group)	1.79 (0.70 to 4.58)	Not relevant	1.90 (0.58 to 6.19)	2.24 (0.75 to 6.84)
CBT individual (over 15 sessions)	0.74 (0.37 to 1.49)	Not relevant	3.99 (2.10 to 7.74)	2.39 (1.13 to 5.08)
CBT group (under 15 sessions)	0.67 (0.30 to 1.46)	Not relevant	2.69 (1.01 to 7.08)	1.69 (0.65 to 4.50)
IPT	0.87 (0.42 to 1.79)	Not relevant	2.30 (1.09 to 4.87)	2.06 (0.92 to 4.61)
Short term PDPT individual	1.05 (0.47 to 2.37)	Not relevant	2.44 (1.04 to 5.75)	1.53 (0.51 to 4.46)
Non-directive counselling	0.47 (0.15 to 1.46)	Not relevant	1.86 (0.44 to 7.54)	3.05 (0.70 to 14.50) (directive counselling effect)
Computerised CBT with support	1.29 (0.58 to 2.95)	Not relevant	2.71 (0.53 to 14.78)	1.01 (0.37 to 2.80)
Computerised CBT without support	1.45 (0.68 to 3.13)	Not relevant	1.60 (0.32 to 8.23)	1.19 (0.44 to 3.23)
Psychoeducational group programme	0.75 (0.29 to 1.90)	Not relevant	1.45 (0.31 to 7.01)	1.41 (0.29 to 7.15)
Physical exercise programme	0.71 (0.35 to 1.46)	Not relevant	2.99 (1.35 to 6.85)	0.84 (0.34 to 2.07)
CBT individual (over 15 sessions) + citalopram (all data borrowed from CBT individual [over 15 sessions] + TCA or amitriptyline)	1.19 (0.38 to 3.77)	3.51 (0.22 to 61.62) (CBT + AD class effect)	1.96 (0.51 to 7.64)	1.00 (0.22 to 4.49)
IPT + citalopram (all data borrowed from IPT + AD)	0.98 (0.22 to 4.50)	0.38 (0.00 to 24.14)	9.05 (2.10 to 39.21)	3.11 (0.56 to 17.12)

Update 2017

Intervention	Mean odds ratios versus baseline (95% credible intervals)			
	Discontinuation versus citalopram	Discontinuation due to side effects in those discontinuing versus SSRIs (respective class effects used)	Response in treatment completers versus pill placebo	Remission in treatment completers versus pill placebo
		(borrowed from short term PDPT + AD class)		
Short term PDPT individual + citalopram (all data borrowed from short term PDPT + AD)	0.78 (0.18 to 3.66)	0.38 (0.00 to 24.14)	3.54 (0.89 to 13.96)	1.61 (0.27 to 10.69)
Physical exercise programme + sertraline	0.77 (0.30 to 1.95)	0.43 (0.02 to 6.90) (Exercise + AD class effect)	2.01 (0.69 to 6.03)	0.92 (0.27 to 3.00)
Clinical management (pill placebo)	1.12 (0.64 to 1.97)	Not relevant	Baseline	Baseline
No treatment following treatment discontinuation (Wait list)	Not relevant	Not relevant	0.37 (0.14 to 1.01)	0.18 (0.06 to 0.58)
Notes: AD: antidepressant; CBT: Cognitive behavioural therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; TCA: tricyclic antidepressant				

Update 2017

Table 290. Results of the NMAs that informed the economic analysis of interventions for a new depressive episode in people with more severe depression: odds ratios versus baseline for each outcome of interest

Intervention	Mean odds ratios versus baseline (95% credible intervals)			
	Discontinuation versus sertraline	Discontinuation due to side effects in those discontinuing versus SSRIs (respective class effects used)	Response in treatment completers versus pill placebo	Remission in treatment completers versus pill placebo
Sertraline	Baseline	Baseline	1.84 (1.15 to 2.93)	1.46 (0.62 to 3.42)
Mirtazapine	0.82 (0.45 to 1.48)	1.44 (0.53 to 3.80)	2.81 (1.66 to 4.80)	0.91 (0.16 to 5.28)
Behavioural activation	1.96 (0.17 to 27.14)	Not relevant	27.36 (1.86 to 362.13)	(probability of remission in responders borrowed from CBT individual [over 15 sessions])
CBT individual (over 15 sessions)	0.69 (0.21 to 2.19)	Not relevant	1.55 (0.21 to 11.06)	1.71 (0.32 to 10.12)

Intervention	Mean odds ratios versus baseline (95% credible intervals)			
	Discontinuation versus sertraline	**Discontinuation due to side effects in those discontinuing versus SSRIs (respective class effects used)**	**Response in treatment completers versus pill placebo**	**Remission in treatment completers versus pill placebo**
CBT group (under 15 sessions)	0.49 (0.11 to 2.24)	Not relevant	3.11 (0.34 to 28.19)	(probability of remission in responders borrowed from CBT individual [over 15 sessions])
Short term PDPT individual	0.36 (0.07 to 1.80)	Not relevant	5.51 (0.93 to 31.28)	1.48 (0.15 to 15.88) (borrowed from IPT)
Non-directive counselling	0.27 (0.04 to 1.97)	Not relevant	4.82 (0.61 to 37.79)	1.48 (0.15 to 15.88) (borrowed from IPT)
Computerised CBT without support	0.65 (0.15 to 2.80)	Not relevant	2.21 (0.23 to 20.88) (borrowed from cognitive bibliotherapy)	0.11 (0.00 to 5.64) (borrowed from computerised CBT with support)
Physical exercise programme	0.17 (0.02 to 1.60)	Not relevant	18.60 (1.26 to 327.01)	0.11 (0.00 to 5.64) (borrowed from computerised CBT with support)
CBT individual (over 15 sessions) + sertraline	0.22 (0.04 to 1.02) (borrowed from CBT individual [over 15 sessions] + SSRI)	0.19 (0.00 to 7.00) (CBT + AD class effect)	6.94 (0.83 to 56.71) (borrowed from CBT individual [over 15 sessions] + SSRI)	4.05 (0.64 to 27.66) (borrowed from CBT individual [over 15 sessions] + imipramine)
Clinical management (pill placebo)	1.15 (0.73 to 1.79)	Not relevant	Baseline	Baseline
No treatment following treatment discontinuation (Wait list)	Not relevant	Not relevant	0.65 (0.05 to 7.41)	0.06 (0.00 to 2.06)

Notes:
AD: antidepressant; CBT: Cognitive behavioural therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; SSRI: Selective serotonin reuptake inhibitor

Update 2017

1

14.2.61 Baseline probabilities

The baseline probabilities of the 4 outcomes of interest were estimated based on published literature and GC expert opinion and were applied in the decision-tree component of the economic model. All relative effects of the other interventions versus the intervention serving as baseline were applied onto the baseline probability in order to obtain the absolute probability of every intervention assessed in the economic analysis for each outcome of interest.

The GC expressed the view that absolute probabilities reported in RCTs included in the NMAs did not reflect probabilities seen under non-interventional conditions and routine clinical practice, and therefore these were not utilised in the economic analysis.

14.2.6.11 Baseline probability of early discontinuation (for any reason)

Burton, Anderson et al. (2012) analysed prescription data from a Scottish primary care database of adults who commenced treatment with an eligible antidepressant between April 2007 and March 2008 across 237 Scottish practices. Eligible antidepressants comprised SSRIs, SNRIs, lofepramine and trazodone. The authors identified 28,027 people who initiated treatment with an eligible antidepressant over this period, of whom 24.6% did not continue treatment beyond 30 days (they discontinued treatment within the first 30 days) and 44.5% did not continue treatment beyond 90 days (they discontinued treatment within the first 90 days). The authors did not report discontinuation rates by level of severity of depression or by specific drug or drug class.

Hansen, Vach et al. (2004) reported rates of discontinuation (defined as people not purchasing antidepressants in the 6 months following first prescription) following analysis of data on 4,860 adult first-time users of antidepressants (regardless of diagnosis) who presented in 174 general practices in Denmark between January 1998 and June 1999. The discontinuation rate was 30.5% for adults prescribed new generation antidepressants, mainly SSRIs (n=4,275) and 56.4% for adults prescribed TCAs (n=585). No information was provided on discontinuation rates in relation to level of severity of symptoms.

Bull, Hunkeler et al. (2002) assessed the rates of discontinuation at 3 and 6 months in 672 adults that were started on an SSRI (fluoxetine or paroxetine) by a psychiatrist or primary care physician for a new or recurrent case of depression between January and September 1998 in the USA. Participants were conducted via a telephone survey. At 3 months, 34% had discontinued their initiated SSRI.

Goethe, Woolley et al. (2007) reported discontinuation data on 406 adults with severe depression who were treated with SSRIs in a secondary care setting (208 as outpatients and 198 as inpatients) in the USA between July 2001 and January 2003. The reported discontinuation rate at 3 months was 24.6%.

Lewis, Marcus et al. (2004) reported rates of early discontinuation among 26,888 adults who filled an SSRI prescription, by analysing data from a large database in the USA. Of these, 61.3% were seen in primary care, 14.9% were treated by psychiatrists and another 23.8% were treated by another medical specialist. Early discontinuation was defined as failure to refill a prescription for any antidepressant medication within 30 days of the end of the first SSRI prescription. The authors reported early discontinuation of 37.1% for adults prescribed an SSRI by primary care providers, 31.8% for those treated by psychiatrists and 41.4% for those treated by other medical specialists. No information was provided on discontinuation rates in relation to level of severity of symptoms.

Olfson, Marcus et al. (2006) analysed data on 829 adults with depression who were initiated on antidepressant treatment, derived from the household component of the Medical

Update 2017

Expenditure Panel Survey conducted in the USA for the years 1996 to 2001. The authors reported rates of discontinuation during the first 30 days of treatment and between 31-90 days of treatment by mental status. In the first 30 days of treatment, discontinuation reached 42.7% in adults with "excellent to good" mental status and 42.0% in adults with "fair or poor" mental status. Between 31-90 days of treatment, discontinuation reached 57.3% in adults with "excellent to good" mental status and 41.1% in adults with "fair or poor" mental status. In total, discontinuation over 90 days reached 75% and 65% in adults with "excellent to good" and those with "fair or poor" mental status, respectively. Discontinuation was lower in people taking SSRIs or SNRIs (40.9% in first 30 days, 48.0% in 31-90 days) compared with other new medications (49.9% in first 30 days, 63.0% in 31-90 days) and TCAs and other old antidepressants (45.2% in first 30 days, 68.2% in 31-90 days). Discontinuation in the first 30 days was lower in adults who had private health insurance (39.9%) compared with those who had public (48.6%) or no (50.6%) insurance. No other information was provided on discontinuation rates in relation to severity of depressive symptoms or type of provider (primary or specialist care).

The GC reviewed the data reported in the studies. The figures of 24.6% and 44.5% for continuation up to 30 and 90 days, respectively, that were reported by Burton, Anderson et al. (2012) are directly relevant to primary care practice in the UK; the figure of 44.5% is likely to include people who took a full first course of treatment but did not continue because of treatment failure (lack of efficacy); therefore the risk of discontinuation of initiated treatment prior to completion of a full course lies between the two figures of 24.6% and 44.5%. It is likely that the figure is relevant to SSRIs, since these are among the most commonly used antidepressants. (Hansen, Vach et al. 2004) reported a discontinuation risk of 30.5% over a period of 6 months for SSRIs prescribed in primary care in Denmark. The USA figures are higher, as Lewis, Marcus et al. (2004) reported a 37.1% discontinuation within 30 days for SSRIs prescribed in primary care, while Olfson, Marcus et al. (2006) reported the highest rates, 75% and 65% over 90 days, in adults with 'excellent to good' and those with 'fair or poor' mental status, respectively. Discontinuation rates were reported to be higher in people treated in primary compared with specialist care.

Following consideration of the data and expert GC opinion, it was decided to use a figure of 37% for early discontinuation of SSRIs in people with less severe depression, and 34% for early discontinuation of SSRIs in people with more severe depression. These figures are within the range of percentages reported by Burton, Anderson et al. (2012) for 30 and 90 days, but lower than the figures reported by Olfson, Marcus et al. (2006) over 90 days. Discontinuation was assumed to be higher in people with less severe depression, based on data reported in Olfson, Marcus et al. (2006) and the GC expert opinion.

The figure of 37% was used as the baseline probability of discontinuation for citalopram, in the economic analysis for people with less severe depression. The figure of 34% was used as the baseline probability of discontinuation for sertraline in the economic analysis for people with more severe depression.

14.2.6.2 Baseline probability of discontinuation due to side effects in those discontinuing treatment early

Discontinuation due to side effects was relevant to cohorts treated with pharmacological treatments or combined treatments with a pharmacological intervention component.

Bull, Hunkeler et al. (2002) reported reasons for drug discontinuation at 3 and 6 months in 672 adults that were started on an SSRI (fluoxetine or paroxetine) by a psychiatrist or primary care physician for a new or recurrent case of depression between January and September 1998 in the USA. Participants were conducted via a telephone survey. Overall, 15% of people who were initiated on a SSRI discontinued due to intolerable side effects over the first 3 months of the study.

Update 2017

Goethe, Woolley et al. (2007) reported discontinuation data on 406 adults with severe depression who were treated with SSRIs in a secondary care setting (208 as outpatients and 198 as inpatients) in the USA between July 2001 and January 2003. Overall, 13% of people who were initiated on an SSRI discontinued due to intolerable side effects over the first 3 months of the study.

The risk of discontinuation due to side effects was considered to be independent of the depressive symptom severity. A risk of 0.15 was therefore applied to people initiated on SSRIs with both less severe and more severe depression. Since the risk of discontinuation with SSRI treatment was estimated to be 37% in people with less severe depression and 34% in people with more severe depression, the estimated risk of discontinuation due to side effects in those discontinuing SSRI treatment was estimated to be 0.41 and 0.44 in people with less severe depression and more severe depression, respectively.

The figure of 0.41 was used as the baseline probability of discontinuation due to side effects in those discontinuing citalopram in the economic analysis for people with less severe depression. The figure of 0.44 was used as the baseline probability of discontinuation due to side effects in those discontinuing sertraline in the economic analysis for people with more severe depression.

14.2.6.38 Baseline probability of response and remission in treatment completers

The only study identified in the literature reporting relevant data by level of depressive symptom severity was conducted by Simon, Goldberg et al. (1999), who reported 12-month outcomes of 948 people with major depression attending primary care services who participated in a multinational, longitudinal study conducted at 15 sites in 14 countries including the UK. All study participants had been assessed at baseline by study researchers using the Composite International Diagnostic Interview (CIDI), the 28-item General Health Questionnaire (GHQ), and the Brief Disability Questionnaire (BDQ) and were classified as having mild, moderate or severe major depression. Participants also underwent assessment by their primary care physicians at baseline; depression or a psychological disorder and a comorbid condition was correctly recognised by physicians in 42% of them. However, no information on follow-up care or treatment received was available for any of the participants. At 12 month follow-up the diagnostic status (ICD-10 depressive disorder) of participants was reported by their baseline symptom severity, stratified according to whether they had been recognised by their physicians at baseline. Recognised and unrecognised groups did not differ significantly in change in diagnostic status from baseline. Results were consistent across study sites.

Table 291 shows the 12-month diagnostic status of people who had been diagnosed with mild, moderate and severe depression at baseline, and who had been recognised by their physician to have a depression or another psychological disorder.

Table 291: Diagnostic status at 12 months of people with major depression that were diagnosed by their physicians at baseline, by baseline severity status, as reported in Simon, Goldberg et al. (1999)

12-month status	Baseline mild depression	Baseline moderate depression	Baseline severe depression
Recovery	79.3%	64.5%	54.9%
Mild depression	6.9%	3.2%	7.8%
Moderate depression	6.9%	19.4%	9.8%
Severe depression	6.9%	12.9%	27.5%
TOTAL	100.0%	100.0%	100.0%

Update 2017

It can be seen that at 12-months probability of recovery is highest for people with mild depression (0.79), lower for people with moderate depression (0.65) and lowest for people with severe depression at baseline (0.55). Based on the data above, it is possible to estimate the probability of improvement from baseline to 12 months for each category of symptom severity, considering improvement as movement to a lower level of severity or recovery. For mild depression the probability of improvement equals that of recovery (0.79); for moderate depression improvement of status is reflected by recovery or a move to mild depression (0.68 in total); and for severe, the probability of improvement is reflected in recovery or reduction of symptoms from severe to mild or moderate (0.73).

These data formed the basis for estimating the 3-month probability of response (as expressed by improvement) and remission at baseline in the economic model for people with less severe depression and those with more severe depression. Although the study reported data on both people recognised by their physicians as having a psychological disorder and those that were not recognised, the economic analysis utilised data on people whose disorder was recognised by their physicians, as the study population of the economic analysis comprises adults with recognised depression initiating treatment. The GC advised that reported data be used to represent the baseline probability of response and remission in those completing clinical management. This was decided as there was no information in the study on the specific treatment received by study participants; the GC considered that a mixture of treatments would have been received, with some people having received more intensive treatment and some others less intensive or no treatment. The GC inspected the available 12-month recovery and improvement data reported for each level of symptom severity and expressed the view that, on balance, they reflect baseline changes in status that are observed under clinical management (GP visits).

As reported in Chapter 13, section 13.2.7, synthesis of remission data from cohort studies following people with depression showed that the probability of remission in people with depression follows a Weibull distribution in which the remission rate is proportional to a power of time. People have a higher probability of remission soon after initiation of the depressive episode, and this probability is reduced over time, as they remain in that episode; the cumulative hazard rate for the Weibull distribution is given by the following mathematical formula:

$$H(t) = \lambda t^{\gamma}$$

where lambda (λ) and gamma (γ) are the scale and shape parameters of the distribution, respectively.

Synthesis of relevant cohort data determined the parameters of the Weibull distribution characterising the probability of remission over time. These parameters, shown in Table 292, were estimated using data from studies on cohorts with depression followed over long periods of time, irrespective of their level of symptom severity.

Table 292: Parameters of the Weibull distribution of the probability of remission over time, in people experiencing a depressive episode

Parameter	Mean	SD	Median	95% Credible intervals
Gamma	0.440	0.026	0.440	0.389 to 0.491
Lambda	1.171	0.085	1.168	1.016 to 1.344

In order to estimate the 3-month probabilities of remission and response in people completing clinical management it was assumed that both followed a Weibull distribution with the same shape parameter gamma across all symptom severity levels that was equal to that estimated from synthesis of cohort studies (Table 292). The lambda parameter for response and remission at each level of severity was estimated from the available 12-month data

Update 2017

(Simon et al., 1999). The estimated 3-month probabilities of response and remission at each symptom severity level as well as the estimated hazard ratios of response and remission at each level of severity versus the 'baseline' remission, estimated from data synthesis, are shown in Table 293.

Table 293: Parameters of the Weibull distribution and 3-month probabilities of response and remission, in people experiencing a depressive episode according to their level of symptom severity

Mean values	'Baseline' remission – based on synthesis of studies	Data based on Simon, Goldberg et al. (1999) for people with major depression recognised by their physician					
Parameter		Mild depression		Moderate depression		Severe depression	
		Resp	Remis	Resp	Remis	Resp	Remis
12-month probability	0.69	0.79	0.79	0.68	0.65	0.73	0.55
Hazard (lambda)	1.17	1.58	1.58	1.14	1.03	1.29	0.79
Hazard ratio vs 'baseline' lambda	1 (reference)	1.34	1.34	0.96	0.88	1.10	0.68
Gamma	0.44						
3-month probability	0.46	0.56	0.56	0.45	0.42	0.49	0.34
Notes: Resp: response; Remis: remission							

The 3-month probabilities of response and remission for people with less severe depression was estimated as an average of respective probabilities estimated for people with mild and moderate depression (0.51 and 0.49, respectively). The 3-month probabilities of response and remission for people with more severe depression were assumed to equal those estimated for people with severe depression (0.49 and 0.34 respectively).

When running the probabilistic analysis, the number of people reaching remission were not allowed to exceed the number of people responding to treatment. In iterations where the probability of remission exceeded the probability of response, the number of people in remission was forced to equal that of people in response (so that all people who responded also remitted in those iterations).

Update 2017

14.2.7.18 Other clinical input parameters

14.2.7.19 Progression of depression in people who responded to acute treatment without reaching remission

People who responded to initial treatment but did not meet criteria for remission at the end of the 12 weeks of treatment were assumed to receive a course of further treatment and either remit or remain in a depressive episode. For the purposes of simplicity, people in this branch of the model were assumed to move to one of the two respective states of the Markov model (remission or depressive episode) at the end of 12 weeks, although in reality this transition would not occur immediately. The probability of moving to the Markov remission state was based on the GC expert opinion, due to lack of relevant data. According to the GC expert opinion, the probability of moving to the Markov remission state in people who had responded to the new treatment but had not reached levels of remission at 12 weeks was 0.60 in less severe depression and 0.30 in more severe depression.

14.2.7.21 Risk of relapse in the Markov component of the economic model

The risk of relapse in people who were in the remission state in the Markov component of the economic model was determined by the time spent in the remission state (one or two years),

the number of previous episodes experienced by each cohort assessed in the analysis, and by the efficacy of relapse preventive treatment, in people who received maintenance treatment.

Baseline risk of relapse

As reported in Chapter 13, section 13.2.6, the risk of relapse in people with depression that is in remission is dependent on time, following a Weibull distribution in which the relapse rate is proportional to a power of time. People have a higher risk of relapse in the early years following remission, and this risk is reduced with every year they remain in remission; the cumulative hazard rate for the Weibull distribution is given by the following mathematical formula:

$$H(t) = \lambda t^{\gamma}$$

where lambda (λ) and gamma (γ) are the scale and shape parameters of the distribution, respectively.

Moreover, there is evidence that the risk of relapse increases with the number of previous episodes.

Synthesis of data from cohort studies following people who remitted from a single (first) episode of depression determined the parameters of the Weibull distribution characterising the baseline risk of relapse after remission of a single episode over time. These parameters are shown in Table 294. Their use in the model allowed estimation of the baseline risk of relapse in people in the remission state according to the time they remained in the state (one or two years).

Table 294: Parameters of the Weibull distribution of risk of relapse over time, in people who are in remission following a single (first) episode

Parameter	Mean	SD	Median	95% Credible intervals
Gamma	0.612	0.057	0.611	0.503 to 0.723
Lambda	0.095	0.010	0.094	0.077 to 0.115

The increase in the risk of relapse for every additional depressive episode was considered by applying the hazard ratio of relapse with every additional episode as estimated by Kessing and Andersen (1999), who reported the results of a case register study that included all hospital admissions with primary affective disorder in Denmark during 1971-1993. A total of 7,925 unipolar patients were included in the study. The authors reported that the risk of relapse increased with every new episode by a mean hazard ratio of 1.15 (95% CI 1.11-1.18). Use of this ratio allowed estimation of the baseline relapse risk for people who, following successful treatment, recovered from their fourth episode.

Risk of relapse associated with interventions aiming at relapse prevention

The effect of relapse preventive treatments in people who completed acute treatment and moved to the remission state in the Markov component of the model was expressed as a hazard ratio versus baseline, and was applied onto the baseline risk of relapse over the first 2 years of the Markov model. The hazard ratios of maintenance treatments versus baseline (clinical management, expressed by pill placebo trial arms) were derived from the NMAs conducted for this guideline to inform the relapse prevention guideline economic models, as described below.

The hazard ratios versus clinical management that were utilised in the Markov component of this economic analysis for cost-effective maintenance treatments were obtained from the

Update 2017

relapse prevention model conducted for this guideline and are presented in Table 295. Hazard ratios of relapse preventive interventions were determined by the acute treatment that led to people's remission, as estimated in Chapter 13, section 13.2.5. The hazard ratios of 4 sessions of psychological interventions received as maintenance treatment were assumed to equal the hazard ratios of maintenance individual cognitive therapy (CT) that was received by people who had remitted following acute CT or maintenance individual CBT and clinical management (drug tapering) in people who had remitted following acute combined treatment, as appropriate, in the guideline relapse prevention economic analysis. The hazard ratio of maintenance group CBT was assumed to equal that of maintenance group CT.

Table 295. Hazard ratios of cost-effective maintenance treatments received by people remitting from a new episode of depression - Results of the NMAs conducted to inform the guideline economic analyses of interventions aiming at relapse prevention in people with depression that is in remission

Intervention	Mean hazard ratio versus pill placebo (95% credible intervals)
People with more severe depression who remitted following acute pharmacological treatment	
Maintenance AD treatment	0.51 (0.46 to 0.56)
MBCT + clinical management (drug tapering)	0.45 (0.33 to 0.59)
People with more severe depression who remitted following acute psychological treatment	
4 sessions of intervention received as acute treatment (assumed to equal effect of maintenance individual CT)	0.72 (0.44 to 1.10)
MBCT	0.91 (0.36 to 1.96)
Group CBT	1.02 (0.36 to 2.32)
People with more severe depression who remitted following acute combined treatment	
Maintenance AD treatment	0.43 (0.27 to 0.64)
4 sessions of psychological intervention received as acute treatment + clinical management (drug tapering)	0.68 (0.44 to 1.00)

Update 2017

In sensitivity analysis, people who remitted across all cohorts were assumed to receive no maintenance treatment and thus to be subject to the (same) baseline risk of relapse.

14.2.7.37 Probability of remission in the Markov component of the economic model

The probability of remission in people who are in the depressive episode state in the Markov component of the economic model was determined by the time spent in the depressive episode state. As discussed in section 14.2.6.3, the probability of remission in people with depression follows a Weibull distribution in which the remission rate is proportional to a power of time. People have a higher annual probability of remission in the early years following initiation of the depressive episode, and this probability is reduced with every year they remain in the episode.

Synthesis of data from cohort studies following people with depression determined the parameters of the Weibull distribution characterising the probability of remission over time, as it has been shown in Table 292. Their use in the model allowed estimation of the risk of remission in people in the depressive episode state according to the time they remained in the state (one or two years).

These parameters were estimated using data from studies on cohorts with depression followed over long periods of time, irrespective of their level of symptom severity.

In order to estimate the Weibull parameters of remission for people with less severe depression and people with more severe depression, data were taken from the study by Simon, Goldberg et al. (1999), details of which are provided in section 14.2.6.3. The probability of remission at 12 months by baseline symptom severity reported in this study was used to estimate lambda parameters for the underlying distribution at each level of symptom severity. The shape parameter gamma that was estimated for recovery from synthesis of cohort studies (reported in Chapter 13, section13.2.7) was assumed to apply across all symptom severity levels. This way a Weibull distribution for recovery was determined for each level of symptom severity; details of the distribution for each level of recovery have been shown in Table 293.

The probability of remission for people with less severe depression in their first and second year in the depressive episode state of the Markov model was estimated as an average of respective probabilities estimated for people with mild and moderate depression using the Weibull parameters shown in Table 293. The probability of remission for people with more severe depression in their first and second year in the depressive episode state of the Markov model was estimated using the Weibull parameters for people with severe depression shown in the same table.

People who entered the Markov component via the depressive state were already in non-remission for 12 weeks and therefore their probability of remission in the first and second year following entrance to the Markov depressive state corresponded to model time points between 12-64 weeks and 64-116 weeks, respectively. This was accounted for in the estimation of probability of remission for this sub-group in the economic analysis.

14.2.7.4 Probability of development of side effects from antidepressant treatment

Update 2017

Treatment with antidepressants is associated with the development of various side effects. These can be serious, including death, attempted suicide or self-harm, falls, fractures, stroke or transient ischaemic attack, epilepsy/seizures, myocardial infarction, hyponatraemia and upper gastrointestinal bleeding (Coupland, Dhiman et al. 2011, Jakobsen, Katakam et al. 2017) or less serious but more common, such as headaches, nausea and other gastrointestinal symptoms, dizziness, agitation, sedation, sexual dysfunction, tremor, sweating, fatigue, and arrhythmia (Anderson, Pace et al. 2012, Jakobsen, Katakam et al. 2017).

Serious side effects from antidepressants are costly to treat and are likely to reduce the HRQoL of people who experience them more significantly compared with less serious side effects. However, they do not occur frequently. Coupland, Dhiman et al. (2011) investigated the association between antidepressant treatment and the risk of several potential adverse outcomes in older people with depression, in a retrospective cohort study that utilised data from 60,746 people aged 65 and over diagnosed as having a new episode of depression, obtained across 570 general practices in the UK between 1996 and 2008. The authors reported that SSRIs were associated with the highest adjusted hazard ratios for falls (1.66, 95%; CIs 1.58 to 1.73) and hyponatraemia (1.52; 95% CIs 1.33 to 1.75) compared with when antidepressants were not being used, while a group of ‘other antidepressants’ defined according to the British National Formulary, which included mirtazapine and venlafaxine, among others, was associated with the highest adjusted hazard ratios for all-cause mortality (1.66; 95% CIs 1.56 to 1.77), attempted suicide or self-harm (5.16; 95% CIs 3.90 to 6.83), stroke/transient ischaemic attack (1.37; 95% CIs 1.22 to 1.55), fracture (1.64; 95% CIs 1.46 to 1.84), and epilepsy/seizures (2.24; 95% CIs 1.60 to 3.15), compared with when antidepressants were not being used. However, for most of these side effects, with the exception of all-cause mortality, the difference in absolute risks between people who received antidepressants and those who did not were small (lower than 1%) with few exceptions: considering the drugs and classes that were included in the guideline economic analysis, for SSRIs, the absolute increase in risk of falls compared with people who did not

take antidepressants was 2.21%; for mirtazapine, the absolute increase in risk of attempted suicide or self-harm compared with people who did not take antidepressants was 1.31%. It is noted that these data were derived from older adults with depression, who are likely to have a higher baseline risk for these events compared with younger populations. Therefore, the absolute increase in risk for any of these events in the study population, between those taking antidepressants and those not taking antidepressants, is expected to be lower than that observed between respective groups in older populations.

Jakobsen, Katakam et al. (2017) conducted a systematic review and meta-analysis to assess the effects (including adverse events) of SSRIs versus placebo, 'active' placebo, or no intervention in adult participants with major depressive disorder. The authors reported that SSRIs significantly increased the risks of serious adverse events (odds ratio 1.37; 95% CI 1.08 to 1.75) corresponding to 31/1000 SSRI participants experiencing a serious adverse event compared with 22/1000 control participants (that is a 0.9% difference).

Anderson, Pace et al. (2012) estimated the prevalence of common side effects such as headaches, nausea or vomiting, agitation sedation and sexual dysfunction associated with treatment with antidepressants, by undertaking a retrospective analysis of data derived from a large USA managed care claims form on 40,017 people aged 13 years and above, of whom 36,400 were adults aged 19 years and above, who were newly diagnosed with depression and were initiated on antidepressant monotherapy between 1998 and 2008. Antidepressant groups included, among others, SSRIs and tetracyclic antidepressants (which, in 99% of cases, were represented by mirtazapine). The authors reported that the most common side effects of those assessed were headaches (5.5 and 6.8/1000 person-months of therapy in adults taking SSRIs and mirtazapine, respectively) followed by nausea (3.6 and 5.5/1000 person-months of therapy in adults taking SSRIs and mirtazapine, respectively). The rate of experiencing at least one of the 5 common side effects considered in the study was 9.7/1000 person-months of therapy in adults taking SSRIs and 13.6/1000 person-months of therapy in adults taking mirtazapine. These translate into 11.7 and 16.3/100 person-years of therapy.

The economic model considered the impact of common side effects on treatment costs and people's HRQoL. A proportion of people receiving SSRIs alone or in combination and those receiving mirtazapine were assumed to be experiencing common side effects at any time over the duration of the model. These proportions equalled 0.117 and 0.163 for SSRIs and mirtazapine, respectively, based on the data reported by Anderson, Pace et al. (2012). No side effects were considered for people receiving non-pharmacological interventions; however, people receiving non-pharmacological interventions are also expected to experience a range of events such as headaches, nausea or vomiting, etc. The study by (Anderson, Pace et al. 2012) was uncontrolled and did not examine the rate of side effects that were attributable to drugs. Therefore, the economic analysis may have overestimated the impact of common side effects from antidepressants relative to other treatments and thus underestimated their relative cost effectiveness.

The economic model did not incorporate the impact of less common but more severe side effects on costs and people's HRQoL, as this would require most complex modelling and detailed data on the course and management of these side effects. However, omission of these severe side effects is not expected to have considerably affected the results of the economic analysis, due to their low incidence in the study population. Nevertheless, omission of less common but severe side effects from the economic analysis may have potentially overestimated the cost effectiveness of pharmacological and combined treatments.

Update 2017

14.2.7.5 Mortality

Depression is associated with an increased risk of mortality relative to the general population. A comprehensive systematic review of 293 studies that assessed the increased risk of people with depression relative to non-depressed individuals, which included

1,813,733 participants (135,007 depressed and 1,678,726 non-depressed) reported a risk ratio of mortality in depressed relative to non-depressed participants of 1.64 (95% CI 1.56 to 1.76). After adjustment for publication bias, the overall risk ratio was reduced to 1.52 (95% CI 1.45 to 1.59) (Cuijpers, Vogelzangs et al. 2014).

The risk of mortality for people with a new episode of depression was not considered in the decision-tree part of the model (12 weeks), because death (mainly due to suicide) is a rare outcome in RCTs of acute treatments for depression, and no substantial differential data on mortality or, specifically, on the risk of suicide between treatments assessed in the economic analysis are available.

In the Markov component of the model, the adjusted risk ratio of mortality in depressed relative to non-depressed participants (Cuijpers, Vogelzangs et al. 2014) was applied onto general mortality statistics for the UK population (ONS 2015), to estimate the absolute annual mortality risk in people experiencing a depressive episode relative to people not experiencing a depressive episode within each cycle of the model. People with a depressive episode were assumed to be at increased mortality risk due to depression only in the years they experienced a depressive episode. The same mortality risk was assumed for both men and women experiencing a relapse, as no gender-specific data were reported in the study. People not experiencing a depressive episode in each model cycle were assumed to carry the mortality risk of the general UK population.

14.2.8 Utility data and estimation of quality adjusted life years (QALYs)

In order to express outcomes in the form of QALYs, the health states of the economic model (remission, response not reaching remission, no response or relapse) need 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.

Update 2017

The systematic review of utility data on depression-related heath states identified 5 studies that reported utility data corresponding to depression-related health states, which were derived from EQ-5D measurements on adults with depression valued by the general UK population (Sapin, Fantino et al. 2004, Kaltenthaler, Brazier et al. 2006, Sobocki, Ekman et al. 2006, Sobocki, Ekman et al. 2007, Mann, Gilbody et al. 2009, Koeser, Donisi et al. 2015). Three of the studies analysed EQ-5D data obtained from adults with depression or common mental health problems participating in RCTs conducted in the UK (Kaltenthaler, Brazier et al. 2006, Mann, Gilbody et al. 2009, Koeser, Donisi et al. 2015). The other two studies analysed naturalistic primary care EQ-5D data from adults with depression in France (Sapin, Fantino et al. 2004) and in Sweden (Sobocki, Ekman et al. 2006, Sobocki, Ekman et al. 2007). All studies reported utility values associated with severity of depression (e.g. mild, moderate or severe) and/or states of depression relating to treatment response (e.g. response, remission, no response) and were thus relevant to the health states considered in economic modelling conducted for this guideline. All studies defined health states using validated measures of depressive symptoms, such as the BDI, the HAMD-17, the PHQ-9, the MADRS and the CGI.

An overview of the study characteristics, the methods used to define health states, and the health-state utility values reported by each of the studies is provided in Table 296.

Table 296: Summary of available EQ-5D derived health-state utility data for depression (UK tariff)

Study	Definition of health states	Health state / severity	N	Mean (SD or 95% CI)
Kaltenthaler, Brazier et al. (2006)	Analysis of EQ-5D and CORE-OM data obtained 62 people with common mental health problems participating in a multi-centre RCT of supervised self-help CBT in the UK (Richards, Barkham et al. 2003). CORE-OM data were first mapped onto the BDI, which was used to categorise people into 3 groups of mild to moderate, moderate to severe and severe depression. BDI cut-off scores used for categorisation were not reported. EQ-5D utility value for no depression obtained from age- and gender-matched normal population in the UK (Kind, Dolan et al. 1998).	No depression Mild to moderate Moderate to severe Severe	NA NR NR NR	0.88 (0.22) 0.78 (0.20) 0.58 (0.31) 0.38 (0.32)
Koeser, Donisi et al. (2015)	Analysis of EQ-5D and HAMD17 data obtained from people with recurrent depression in full or partial remission participating in a RCT of MBCT in the UK (N=123) (Kuyken, Byford et al. 2008). Definition of health states by HAMD scores: remission ≤ 7; response 8-14; no response ≤ 15	Remission Response No response	NR NR NR	0.80 (0.02) 0.62 (0.04) 0.48 (0.05)
Mann, Gilbody et al. (2009)	Analysis of EQ-5D and PHQ-9 data collected from 114 people with depression participating in a cluster RCT of collaborative care across 19 UK primary care practices based in urban and rural communities (Richards, Lovell et al. 2008). Definition of health states by PHQ-9 score: mild 5-9; moderate 10-14; moderately severe 15-19; severe 20-27	Mild Moderate Moderate to severe Severe	10 24 39 35	0.65 (0.23) 0.66 (0.21) 0.56 (0.27) 0.34 (0.29)
Sapin, Fantino et al. (2004)	Analysis of EQ-5D and MADRS data collected from 250 people with major depression recruited from 95 French primary care practices for inclusion in an 8-week follow-up cohort. Definition of health states by MADRS score: remission MADRS ≤ 12; response at least 50% reduction in the MADRS baseline score over 8 weeks. Baseline mean MADRS score 32.7 (SD 7.7)	Response – remission Response – no remission No response Baseline	144 34 46 250	0.85 (0.13) 0.72 (0.20) 0.58 (0.28) 0.33 (0.25)
Sobocki, Ekman et al. (2006) **Sobocki, Ekman et al. (2007)**	Analysis of EQ-5D and CGI-S and CGI-I data collected from 447 adults with depression enrolled in a naturalistic longitudinal observational 6-month study conducted in 56 primary care practices in 5 regions of Sweden. People who started a new or changed antidepressant treatment were eligible for inclusion. Definition of health states by CGI-S score: mild 2-3; moderate 4; severe 5-7; remission 'much or very much improved' score (1-2) combined with clinical judgement	Mild Moderate Severe Remission No remission	110 268 69 207 191	0.60 (0.54 to 0.65) 0.46 (0.30 to 0.48) 0.27 (0.21 to 0.34) 0.81 (0.77 to 0.83) 0.57 (0.52 to 0.60)

Notes:
CI: confidence intervals; N: number of participants who provided ratings on the EQ-5D; NR: not reported; SD: standard deviation

Update 2017

All reported utility data comply with the NICE criteria on selection of utility data for use in NICE economic evaluations (NICE 2013). The data from Kaltenthaler, Brazier et al. (2006) were derived following mapping of CORE-OM data onto BDI data; however, the BDI cut-off scores used to determine the health states by depressive symptom severity were not reported, and therefore it is not clear the exact level of symptom severity the resulting utility scores correspond to. All other studies provided details on the scale cut-off scores used to determine the depression-related health states by severity or by response to treatment.

The economic analysis utilised a combination of data from (Sapin, Fantino et al. 2004) and (Sobocki, Ekman et al. 2006, Sobocki, Ekman et al. 2007) for the states of acute treatment, corresponding to the decision-tree component of the model. This was decided because these two studies provided data for all states included in the model, i.e. less or more severe depression at initiation of treatment or following a relapse, remission, response not reaching remission, and no or inadequate response, and were based on larger study samples compared with the other studies providing utility data. It is noted though, that remission in (Sobocki, Ekman et al. 2006, Sobocki, Ekman et al. 2007) was defined as an improved or very much improved score on the CGI-Improvement scale, combined with a clinical judgement by the treating doctor of being in full remission. It is acknowledged that this definition of remission may actually include response to treatment not reaching full remission.

For less severe depression the utility value corresponding to mild depression (0.60) was used, because the study population with less severe depression includes populations with sub-threshold depression and also populations reaching moderate depression, so on average, their utility was considered to correspond to the reported value of mild depression. For more severe depression, a weighted average of the utility of moderate and severe depression of 0.42 was used (values for both states obtained from (Sobocki, Ekman et al. 2006, Sobocki, Ekman et al. 2007)). For people reaching remission and those responding without reaching remission after acute treatment (i.e. at the end of the decision-tree component of the model) the reported values of 0.85 and 0.72 from Sapin, Fantino et al. (2004) were used, respectively. People with no or inadequate response to treatment were assumed to remain in the same state of less severe (0.60) or more severe (0.42) depression.

Update 2017

For the Markov component of the model, the slightly more conservative value of 0.81, reported by Sobocki, Ekman et al. (2006) and (Sobocki, Ekman et al. 2007), rather than the value of 0.85, reported by Sapin, Fantino et al. (2004), was used for people in remission, to reflect the fact that some people may not be in full remission for the whole model cycle, but may experience some symptoms which, nevertheless, are not adequate to indicate relapse. The values of 0.60 and 0.42 were used for people in the depressive less severe and more severe states, respectively, of the Markov component of the model.

In sensitivity analysis, the values of 0.80 for remission and 0.62 for response not reaching remission reported in Koeser, Donisi et al. (2015) were tested. Moreover, in another scenario, the values of 0.65 and 0.56, reported by Mann, Gilbody et al. (2009) for mild and moderate-to-severe depression were attached to the states of less severe and more severe depression, respectively.

Changes in utility between baseline and endpoint of the decision-tree part of the model were assumed to occur linearly.

According to the GC expert opinion, an average depressive episode lasts 6 months. This estimate is supported by data from a prospective study on 250 adults with a newly originated (first or recurrent) major depressive episode, drawn from a prospective epidemiological Dutch survey on 7,046 people in the general population (Spijker, de Graaf et al. 2002). According to this study, the mean duration of a recurrent episode was 6.1 months (95% CI 4.7-7.5). The economic model assumed that people in the Markov component of the model experiencing a depressive episode that resolved in the next year (i.e. people who spent only

a year in the depressive episode and then moved to the remission state in the next cycle), experienced a reduction in their HRQoL for 6 months out of the 12 months of the cycle they remained in the 'depressive' state. Thus, people relapsing to depressive episodes that lasted only for one year were assumed to have the utility of remission for 6 months and the utility of depression (mild or moderate) for another 6 months. However, people whose depressive episode was expected to last for 2 cycles (years) or more, were attached the utility of depression over the number of years (1 or 2) they remained in the depressive episode except their final year in the episode, in which they were assumed to have the utility of depression for 6 months and the utility of remission for another 6 months.

Side effects from medication are expected to result in a reduction in utility scores of adults with depression. (Sullivan, Valuck et al. 2004) applied regression analysis on EQ-5D data (UK tariffs) obtained from participants in the 2000 national USA Medical Expenditure Panel Survey to derive age-adjusted utility values for health states associated with depression and with side effects of antidepressants. Health states were defined based on descriptions in the International Classification of Diseases (9th Edition) (ICD-9) and the Clinical Classification Categories (CCC) (clinically homogenous groupings of ICD-9 codes derived by the Agency for Healthcare Research and Quality).

Table 297 shows the health states determined by Sullivan, Valuck et al. (2004) and the corresponding utility values obtained from regression analysis of EQ-5D data. The mean utility decrements due to side effects from antidepressants ranged from -0.044 (diarrhoea) to -0.129 (excitation, insomnia and anxiety), with a mean decrement of -0.087. This mean utility decrement was used in the economic model for people who discontinued treatment due to intolerable side effects, as no specific information on the type and frequency of side effects that led to discontinuation was available across RCTs; it was applied over 5 weeks, based on the GC advice on the duration of reduction in HRQoL due to intolerable side effects. This utility decrement was also applied to the proportion of people who completed antidepressant treatment and experienced tolerable side effects, over the whole period of antidepressant treatment, i.e. over 12 weeks (acute antidepressant treatment) and the following 2 years (only in those receiving maintenance antidepressant treatment).

Update 2017

Table 297: Summary of EQ-5D derived health-state utility data for side effects from antidepressants (UK tariff)

Study	Definition of health states	Health state	Mean (95% CI)
(Sullivan, Valuck et al. 2004)	Censored least absolute deviations (CLAD) regression analysis of EQ-5D data from the 2000 national US Medical Expenditure Panel Survey (MEPS) [http://meps.ahrq.gov/mepsweb/] Definitions of health states Gastrointestinal symptoms (GI): average Diarrhoea: clinical classification categories (CCC) - Agency for Healthcare Research and Quality): 144 regional enteritis Dyspepsia: CCC 138 oesophageal disorders Nausea & constipation: assumed average of GI Sexual: ICD-9 302 sexual disorders Excitation: average Insomnia: assumed equal to anxiety Anxiety: CCC 072 anxiety, somatoform, dissociative disorders Headache: CCC 084 headache Drowsiness & other: assumed average of all side effects Untreated depression ICD-9 311 depressive disorder; CLAD 25% Treated depression: ICD-9 311 depressive disorder; CLAD 75%; baseline utility estimate (not a decrement)	GI symptoms Diarrhoea Dyspepsia Nausea Constipation Sexual Excitation Insomnia Anxiety Headache Drowsiness Other Untreated depression Treated depression	-0.065 (-0.082 to -0.049) -0.044 (-0.056 to -0.034) -0.086 (-0.109 to -0.065) -0.065 (-0.082 to -0.049) -0.065 (-0.082 to -0.049) -0.049 (-0.062 to -0.037) -0.129 (-0.162 to -0.098) -0.129 (-0.162 to -0.098) -0.129 (-0.162 to -0.098) -0.115 (-0.144 to -0.087) -0.085 (-0.107 to -0.065) -0.085 (-0.107 to -0.065) -0.268 (-0.341 to -0.205) 0.848 (0.514 to 0.971)

Update 2017

14.2.9 Intervention resource use and costs

Intervention costs were estimated by combining resource use associated with each intervention with appropriate unit costs (drug acquisition costs and healthcare professional unit costs).

14.2.9.1 Pharmacological interventions

Pharmacological intervention costs consisted of drug acquisition and GP visit costs. In addition to citalopram, sertraline and mirtazapine, the model also considered clinical management (reflected in the pill placebo arms of the RCTs included in the NMAs that informed this economic analysis), which comprised GP visits only.

The average daily dosage for each drug was determined according to optimal clinical practice (BNF 2016), following confirmation by the GC expert opinion to reflect routine clinical practice in the NHS, and was consistent with dosages reported in the RCTs that were included in the RCTs of pharmacological interventions included in the NMA.

Titration was not explicitly considered in the model; however, in each cohort different percentages of people were allowed to receive different drug daily doses to reflect that some people require titration to a higher dose to achieve optimal intervention effects.

Acute pharmacological treatment was administered over 12 weeks. At the end of this period, people who achieved remission either received maintenance pharmacological treatment with the same drug, or received MBCT combined with gradual discontinuation (tapering) of the drug, which was modelled as a linear reduction of the drug acquisition cost (from optimal dose to zero) at the beginning of maintenance treatment and over a period of one month, according to routine clinical practice, as advised by the GC.

Provision of acute pharmacological treatment involved 4 GP visits. Four GP visits were also assumed for people under clinical management (pill placebo). These resource use estimates were based on the GC expert advice; they represent UK optimal routine clinical practice but may be lower than some of the descriptions of medical resource use in pharmacological trial protocols, where resource use is more intensive than clinical practice.

The drug acquisition costs and the GP unit cost were taken from national sources (Curtis and Burns 2016, NHS 2017). The reported GP unit cost included remuneration, direct care staff costs and other practice expenses, practice capital costs and qualification costs. The latter represented the investment costs of pre-registration and postgraduate medical education, annuitised over the expected working life of a GP; ongoing training costs were not considered due to lack of available information. The unit cost per patient contact was estimated taking into account the GPs' working time as well as the ratio of direct (surgeries, clinics, telephone consultations & home visits) to indirect (referral letters, arranging admissions) patient care, and time spent on general administration.

Intervention costs of acute pharmacological treatment and clinical management are shown in Table 298.

Table 298: Intervention costs of pharmacological interventions for the acute treatment of adults with a new episode of depression considered in the guideline economic analysis (2016 prices)

Drug	Mean daily dosage	Drug acquisition cost[1]	12-week drug cost	Total intervention cost (drug and GP[2]) – acute treatment
Citalopram	50% 20mg 50% 40mg	20mg, 28 tab, £0.83 40mg, 28 tab, £1.01	£2.43	£146.73

Update 2017

Drug	Mean daily dosage	Drug acquisition cost[1]	12-week drug cost	Total intervention cost (drug and GP[2]) – acute treatment
Sertraline	50% 50mg 50% 100mg	50mg, 28 tab, £1.13 100mg, 28 tab, £1.26	£3.59	£147.59
Mirtazapine	50% 30mg 50% 45mg	30mg, 28 tab, £1.19 45mg, 28 tab, £1.50	£4.23	£148.23
Pill placebo (clinical management)	Non-applicable	Non-applicable	Non-applicable	£144.00

Notes:
1 NHS (2017)
2 GP cost includes 4 visits for active acute pharmacological treatment and 4 visits for clinical management; GP unit cost £36 per patient contact lasting 9.22 minutes (Curtis and Burns 2016)

14.2.9.2 Psychological interventions

Resource use estimates of each psychological therapy in terms of number and duration of sessions, mode of delivery and number of therapists and participants in the case of group interventions were determined by resource use data described in respective RCTs that were included in the NMA that informed the economic analysis, modified by the GC to represent routine clinical practice in the UK. High intensity individual psychological interventions were assumed to be delivered by an Agenda for Change (AfC) band 7 clinical psychologist. The other psychological interventions (self-help with support, group CBT, Coping with Depression group course) were assumed to be delivered by an AfC band 5 psychological well-being practitioner (PWP). These assumptions were based on the GC expert advice regarding the delivery of psychological interventions in routine clinical practice, although it was acknowledged that there may variation in the types of therapists delivering psychological interventions across different settings in the UK. For this reason and in order to explore the impact of therapist unit cost on the results of the economic analysis, in deterministic sensitivity analysis, high-intensity psychological interventions were estimated to be delivered by a band 5 PWP and group psychological interventions (group CBT and Coping with Depression group course) were assumed to be delivered by band 7 clinical psychologists. Further to these scenarios, in deterministic sensitivity analysis the number of counselling sessions was reduced to 8 (from 16, which was the number of counselling sessions in base-case analysis), to reflect the fact that some RCTs assessed a lower number of sessions for counselling.

Therapist unit costs were estimated using a combination of data derived from national sources (British-Association-for-Behavioural-&-Cognitive-Psychotherapies 2016, Curtis and Burns 2016, National-College-for-Teaching-and-Leadership 2016) and included wages/salary, salary on-costs, capital and other overheads, qualification costs and the cost of monthly supervision. In estimating the unit cost of each type of therapist per hour of client contact, the ratio of direct (face-to-face) to indirect time (reflecting time for preparation of therapeutic sessions and other administrative tasks) of the therapist was also taken into account.

The unit cost of a band 7 clinical psychologist was estimated to be £97 per hour of direct contact with the client. Details on the method of estimation of the unit cost of a clinical psychologist band 7 are provided in Chapter 13, section 13.2.11.2. An overview of the cost elements that were taken into account in this estimation is shown in Table 299.

Update 2017

Table 299: Unit cost of clinical psychologist band 7 (2016 prices)

Cost element	Unit cost (annual)	Source
Wages – salary	£38,173	Curtis and Burns (2016); unit cost of MBCT therapist (Agenda for Change band 7)
Salary on-costs	£9,500	
Overheads – staff	£11,680	
Overheads - non-staff	£18,211	
Capital overheads	£4,583	
Qualifications	£9,673	Based on a mean clinical psychologist training cost estimate of £159,420 (National-College-for-Teaching-and-Leadership 2016) and a working life of 25 years
Supervision	£306	Based on the unit cost of an Agenda for Change band 8a clinical psychologist (Curtis and Burns 2016) providing 1.5 hour of supervision per month, delivered in groups of 4 participants (British-Association-for-Behavioural-&-Cognitive-Psychotherapies 2016) and expert advice); qualification costs included, assuming a working life of 25 years (National-College-for-Teaching-and-Leadership 2016).
SUM of unit costs	**£92,126**	
Working time	42.4 weeks /year 37.5 hours /week (1,590 hours)	Curtis and Burns (2016)
Total cost per hour	**£58**	
Ratio of direct to indirect time*	1:0.67	Curtis and Burns (2016); estimate supported by GC expert opinion and a review of respective ratios reported in the literature for clinical psychologists and other therapists delivering psychological interventions
Estimated cost per hour of direct contact	**£97**	

Note:
* ratio of face-to-face time to time for preparation and other administrative tasks

Update 2017

The unit cost of band 5 PWP was estimated to be £42 per hour of direct contact with the client. An overview of the cost elements that were taken into account in this estimation is shown in Table 300.

Table 300: Unit cost of psychological well-being practitioner band 5 (2016 prices)

Cost element	Unit cost (annual)	Source
Wages – salary	£23,319	Curtis and Burns (2016); unit cost for community-based scientific and professional staff band 5
Salary on-costs	£5,370	
Overheads – staff	£7,029	
Overheads - non-staff	£10,960	
Capital overheads	£4,583	
Qualifications	£601	Based on a mean training cost estimate of £5,000 (GC expert advice) and a working life of 10 years

Cost element	Unit cost (annual)	Source
Supervision	£1,391	Based on the unit cost per hour of an Agenda for Change band 7 clinical psychologist (as estimated in Table 299) providing 2 hours of individual supervision per month
SUM of unit costs	**£53,253**	
Working time	42.7 weeks /year 37.5 hours /week (1,603 hours)	Curtis and Burns (2016)
Total cost per hour	**£33**	
Ratio of direct to indirect time*	1:0.25	assumption based on GC expert opinion
Estimated cost per hour of direct contact	**£42**	
Note: * ratio of face-to-face time to time for preparation and other administrative tasks		

In addition to therapists' time, the intervention costs of all psychological therapies included an initial GP visit for referral to psychological services.

Moreover, the intervention costs of computerised self-help therapies included the cost of the provider of digital mental health programmes and related equipment required for their delivery (personal computers [PCs] and capital overheads). The cost of provision of a computerised CBT programme per client by the main provider of digital mental health programmes comprises a fixed fee of £36.20, which is independent of the number of sessions attended (GC expert advice). The annual costs of hardware and capital overheads (space around the PC) were based on reported estimates made for the economic analysis undertaken to inform the NICE Technology Appraisal on computerised CBT for depression and anxiety (Kaltenthaler, Brazier et al. 2006) and equal £169 and £1,120, respectively (in 2016 prices). Kaltenthaler, Brazier et al. (2006) estimated that one PC can serve around 100 people with mental disorders treated with computerised programmes per year. Assuming that a PC is used under full capacity (that is, it serves no less than 100 people annually, considering that it is available for use not only by people with depression, but also by people with other mental health conditions), the annual cost of hardware and capital overheads was divided by 100 users, leading to a hardware and capital overheads cost per user of £13. It must be noted that if users of such programmes can access them from home or a public library, then the cost of hardware and capital overheads to the NHS is zero.

Details on resource use and total costs of psychological interventions (or elements of combined interventions) are provided in Table 301.

Table 301: Intervention costs of psychological therapies for adults with a new episode of depression considered in the guideline economic analysis (2016 prices)

Intervention	Resource use details	Total intervention cost per person[1]
Computerised CBT with support	1 session of 45 minutes and 5 sessions of 20 minutes = 2.42 therapist hours per service user (band 5 PWP); fixed cost of provider of digital mental health programmes is £36.20 per person (GC information); cost of hardware & capital overheads £13 per person (2016 price, based on Kaltenthaler, Brazier et al. (2006)	£150 + £36
Computerised CBT without support	Fixed cost of provider of digital mental health programmes is £36.20 per person (GC	£49 + £36

Update 2017

Intervention	Resource use details	Total intervention cost per person[1]
	information); cost of hardware & capital overheads £13 per person (2016 price, based on (Kaltenthaler, Brazier et al. 2006)	
BA	16 sessions x 1 hour each = 16 therapist hours per service user (band 7 clinical psychologist)	£1,545 + £36
Coping with Depression Course (group)	12 sessions x 2 hours each; 2 therapists (band 5 PWPs) and 12 participants per group = 48 therapist hours per group and 4 therapist hours per service user (band 5 PWP)	£166 + £36
CBT individual (over 15 sessions)	16 sessions x 1 hour each = 16 therapist hours per service user (band 7 clinical psychologist)	£1,545 + £36
CBT group (under 15 sessions)	9 sessions x 90 minutes each; 2 therapists (band 5 PWPs) and 12 participants per group = 27 therapist hours per group and 2.25 therapist hours per service user (band 5 PWP)	£93 + £36
Psychoeducational group programme	9 sessions x 90 minutes each; 2 therapists (band 5 PWPs) and 12 participants per group = 27 therapist hours per group and 2.25 therapist hours per service user (band 5 PWP)	£93 + £36
IPT	16 sessions x 1 hour each = 16 therapist hours per service user (band 7 clinical psychologist)	£1,545 + £36
Short term PDPT individual	16 sessions x 1 hour each = 16 therapist hours per service user (band 7 clinical psychologist)	£1,545 + £36
Counselling	16 sessions x 1 hour each = 16 therapist hours per service user (band 7 clinical psychologist)	£1,545 + £36

Notes:
1 cost of psychological intervention plus 1 GP referral visit, at a GP unit cost £36 per patient contact lasting 9.22 minutes (Curtis and Burns 2016); cost of psychological intervention based on resource use combined with unit cost of the appropriate level of therapist, estimated as described in Table 299 and Table 300.
BA: behavioural activation; CBT: cognitive behavioural therapy; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy; PWP: psychological well-being practitioner

Update 2017

14.2.9.3 Physical treatment (physical exercise programme)

Resource use estimates of the physical exercise programme were estimated based on resource use data described in respective RCTs that were included in the guideline NMA that informed the economic analysis, modified by the GC to represent routine clinical practice in the UK. Physical exercise sessions were assumed to be delivered by an AfC band 5 practitioner, with a unit cost equivalent to that of PWP. The PWP unit cost was estimated to be £42 per hour of direct contact with the client. An overview of the cost elements that were taken into account in this estimation is shown in Table 300.

In addition to the PWP's time, the intervention cost of a physical exercise programme included an initial GP visit for referral to exercise sessions. Details on the estimation of the intervention cost of the physical exercise programme are shown in Table 302.

Table 302: Intervention cost of a physical exercise programme for adults with a new episode of depression considered in the guideline economic analysis (2016 prices)

Intervention	Resource use details	Total intervention cost per person[1]
Physical exercise programme	2 weekly group sessions for 5 weeks and 1 weekly group session for another 5 weeks, lasting 45 minutes each; 1 practitioner equivalent, in terms of unit cost, to PWP therapist and 8 participants per group = 11.3 therapist hours per group and 1.4 therapist hours per service user	£58 + £36

Notes:
1 cost of physical exercise programme plus 1 GP visit, at a GP unit cost £36 per patient contact lasting 9.22 minutes (Curtis and Burns 2016); cost of physical exercise programme based on resource use combined with the unit cost of PWP, estimated at £42 per hour of direct client contact as described in Table 300.
PWP: psychological well-being practitioner

14.2.9.4 Combined pharmacological and psychological interventions

The intervention cost of combined interventions was estimated as the sum of the intervention costs of the individual treatment components.

In cohorts receiving combination treatment of pharmacological and psychological interventions or a physical exercise programme, no extra GP visits were added in the psychological intervention or exercise programme cost, since people were already receiving GP care as part of their antidepressant treatment.

Update 2017

14.2.9.5 Interventions received as maintenance treatments aiming at preventing relapses

People who remitted following successful acute treatment moved on to an appropriate relapse preventive intervention, the cost of which was based on the resource use estimates made to inform the guideline economic modelling of interventions for relapse prevention that is described in Chapter 13, section 13.2.11.

An overview of the resource use and cost estimates of relapse preventive interventions used by the cohorts who remitted following successful treatment of a new depressive episode are provided in Table 303.

Table 303: Intervention costs of maintenance treatments considered in the guideline economic analysis on relapse prevention (2016 prices)

Maintenance treatment	Resource use	Total cost
Citalopram	50% of people receiving 20mg/day and the other 50% 40mg/day plus 6 GP visits in the 1st year and 3 GP visits in the 2nd year, plus a visit during tapering	£383
Sertraline	50% of people receiving 50mg/day and the other 50% 100mg/day plus 6 GP visits in the 1st year and 3 GP visits in the 2nd year, plus a visit during tapering	£391
Mirtazapine	50% receiving 30mg/day and the other 50% 45mg/day plus 6 GP visits in the 1st year and 3 GP visits in the 2nd year, plus a visit during tapering	£396
Clinical management - drug tapering	3 GP visits in the first year plus 1 extra GP visit for drug tapering plus linear reduction of the drug dosage over a month; 1 GP visit in the second year	£180-£181 depending on drug
4 sessions of individual	4 individual sessions lasting 1 hour each = 4 therapist hours per service user, plus 2 GP visits	£386 + £72

Maintenance treatment	Resource use	Total cost
psychological therapy		
MBCT	8 group sessions + 4 group booster sessions lasting 2 hours each; 1 therapist and 12 participants per group = 24 therapist hours per group and 2 therapist hours per service user, plus 2 GP visits	£193 + £72
Group CBT	4 group sessions lasting 1.5 hours each; 2 therapists and 12 participants per group = 12 therapist hours per group and 1 therapist hour per service user, plus 2 GP visits	£42 + £72
Clinical management follow-up [no active relapse prevention treatment]	3 GP visits in the first year and 1 GP visit in the second year	£144

Notes:
Unit costs: GP unit cost £36 per patient contact lasting 9.22 minutes (Curtis and Burns 2016); all psychological interventions provided by clinical psychologist band 7, at a unit cost of £97 per hour of direct client contact (Table 299).
CT: cognitive therapy; MBCT: mindfulness-based cognitive therapy

14.2.10 Other healthcare costs considered in the economic analysis

14.2.10.1 Healthcare costs associated with the Markov states of remission and depressive episode

Update 2017

The costs of the states of remission and depressive episode in the Markov component of the economic model were estimated using primarily data from (Byford, Barrett et al. 2011). This was a naturalistic, longitudinal study that aimed to estimate the health service use and costs associated with non-remission in people with depression using data from a large primary care UK general practice research database between 2001 and 2006. The study analysed 12-month healthcare resource use data on 88,935 adults with depression and in receipt of at least two antidepressant prescriptions (for amitriptyline, citalopram, escitalopram, fluoxetine, paroxetine, sertraline or venlafaxine) in the first 3 months after the index prescription. The study provided data on resource relating to medication (antidepressant use and concomitant medication such as anxiolytics, hypnotics, mood stabilizers and neuroleptics), GP contacts, psychological therapy, psychiatrist and other specialist contacts, inpatient stays and accident and emergency attendances. Data were reported separately for people who remitted within 12 months, and those who did not remit.

The study provided cost data for the subgroup of study participants with severe depression. Using the cost figures reported in the paper and the numbers of people in each remission status and symptom severity level it was possible to estimate costs for people with non-severe (mild or moderate) depression. The cost figures corresponding to each remission status and level of symptom severity are shown in Table 304.

Table 304: Healthcare costs of people with depression who remitted within 12 months and people who did not remit within 12 months from index prescription, by symptom severity status, participating in the study by Byford, Barrett et al. (2011)

Remission status	Cost and N in each category		
	All levels of symptom severity N = 88,935 (reported costs)	Severe depression N = 8,106 (reported costs)	Mild or moderate depression N = 80,829 (estimated costs)
People who remitted within 12 months	£656 (N=53,654)	£749 (N=4,423)	£648 (N= 49,231)
People who did not remit within 12 months	£973 (N=35,281)	£1,037 (N=3,683)	£966 (N=31,598)

Costs for severe depression could be potentially attached to states experienced by people with more severe depression in the economic model, while costs for mild or moderate depression could be potentially attached to states experienced by people with less severe depression. However, it can be seen that the mean healthcare costs of people with mild or moderate depression were very similar (only 1% lower) to the respective mean healthcare costs of all participants in the study. Mean costs of people with severe depression were somewhat higher than the mean respective costs of the total study sample (7% higher for people who did not remit and 14% higher for people who remitted). These differences in costs according to symptom severity were not considered to have a substantial impact on the model results. Moreover, people with severe depression in the study may have more severe symptoms than people with more severe depression in the economic analysis. Therefore, it was decided to use the mean total costs reported in the study for the whole study sample (regardless of symptom severity) as the basis for estimation of healthcare costs for people with both less severe and more severe depression. These costs were tested in sensitivity analysis.

Healthcare resource use and cost data reported for the whole study sample in (Byford, Barrett et al. 2011) were modified following GC advice and attached to the health states of the Markov component of the economic model: data on people in a depressive episode who remitted within 12 months in the study were attached onto people in the depressive state of the model if they moved to the remission state (or were expected to remit) in the following year. Resource use and cost data on people who did not remit within 12 months in the naturalistic study were used as the basis for estimating healthcare costs incurred by people who remained (or were expected to remain) in the depressive episode state in the next cycle of the model. Costs incurred after remission was achieved in the naturalistic study were used to estimate annual healthcare costs associated with the remission state of the model. In people that experienced remission whilst being in the Markov component of the model (i.e. not those entering the Markov component in the remission state), an annual cost of maintenance drug treatment plus the cost of 3 GP visits was added to this figure for the first year of remission only, to reflect optimal maintenance antidepressant therapy after remission was achieved, as discussed in Chapter 13, section 13.2.12.

Following GC advice, some of the resource use and drug acquisition cost data reported in the paper were modified, to reflect current clinical practice and the fact that some drugs are now available off patent. Some cost data were sought from other sources. Where detailed resource use data were provided, these were combined with appropriate 2016 unit costs; where only cost figures were available, these have been uplifted to 2016 prices using the hospital & community health services (HCHS) index (Curtis and Burns 2016), so that all costs in the guideline economic analysis reflect 2016 prices.

Details on the methods used to modify and update the resource use and unit costs reported in Byford, Barrett et al. (2011) in order to estimate costs associated with the 2 states of the Markov model component are provided in Chapter 13, section 13.2.12. The healthcare costs associated with each health state in the Markov component of the guideline economic model of treatments for new episodes of depression are presented in Table 305.

Update 2017

Table 305: Annual healthcare costs associated with the states of remission and depressive episode in the guideline economic analysis (2016 prices)

Health state	Cost	Comments
Depressive episode – people remaining (or expected to remain) for longer than one model cycle	£1,483	Includes costs of antidepressants, concomitant medication, GP visits or phone calls, psychological therapy contacts, psychiatrist or other specialist contacts, hospitalisations, and accident and emergency attendances. Costs estimated by multiplying relevant resource use for non-remitters and remitters reported in Byford, Barrett et al. (2011) with appropriate national unit costs for 2016 (Curtis and Burns 2016, Department-of-Health 2016). Treatment costs estimated by published sources of relevant resource use and costs (Radhakrishnan, Hammond et al. 2013, NHS-England 2016). All costs expressed in 2016 prices using the hospital & community health services inflation index (Curtis and Burns 2016) and the estimated net ingredient cost per antidepressant or concomitant medication prescription item ratio for 2015:2006, estimated using national data (NHS-The-Information-Centre 2007, Prescribing & Medicines Team 2016). (Details provided in Chapter 13, Table 278.)
Depressive episode – people moving (or expected to move) to the remission state in the next model cycle	£1,079	
Remission	£493	3-month healthcare cost of people having achieved remission obtained from graphs published by (Byford, Barrett et al. 2011), read using digital software (http://www.digitizeit.de), extrapolated to 12 months and uplifted to 2016 prices using the HCHS inflation index (Curtis and Burns 2016).
Maintenance antidepressant therapy – 1st year extra cost	**£141**	Additional cost reflecting optimal duration of maintenance antidepressant therapy following remission, comprising of an annual antidepressant drug cost equal to that estimated for remitters and 3 GP contacts at the GP unit cost of £36 per patient contact lasting 9.22 minutes for 2016 (Curtis and Burns 2016). This was considered only in people experiencing a remission while being in the Markov model, not in those entering the Markov model in the remission state; the latter received an active relapse preventive intervention or no relapse preventive intervention.

Update 2017

14.2.10.2 Treatment costs in people discontinuing treatment early in the decision-tree component of the model

People who discontinued treatment early consumed part of the acute intervention resources: people who discontinued pharmacological treatment incurred the cost of 1 GP visit and 1 pack of drugs; people who discontinued a high intensity individual psychological therapy incurred the cost of 25% of the visits (i.e. 4 visits) plus the initial GP visit; people who discontinued computerised CBT incurred the cost of the initial GP visit, the full fixed cost of the provider of the programme plus the cost of 2 of the therapist contacts (if they attended a therapist supported programme). People under clinical management who discontinued treatment incurred the cost of 1 GP visit. People who discontinued a group psychological therapy or a physical exercise programme were assumed to incur the full cost of therapy, since participants in a group intervention are not replaced in the group if they discontinue and therefore the full cost of therapy per participant is incurred, whether the participant attends the full course or not.

Those who switched to a mixture of available treatments were assumed to incur a treatment cost over 8 of the 12 weeks of the decision-tree. This cost was estimated as a proportion (8/52) of the annual cost of a depressive episode (for people remaining in depression for longer than one model cycle) that was estimated for the Markov component of the model, which equalled £228.

The cost of no treatment over 8 weeks was assumed to be zero; over this period people receiving no treatment were assumed to incur no depression-specific costs. However, those who entered the depressive state of the Markov model were assumed to re-start receiving depression-related care and incur the cost associated with the depressive Markov state.

14.2.10.3 Cost of management of intolerable or tolerable common side effects from antidepressant treatment

People who discontinued antidepressant or combined treatment due to intolerable side effects were assumed to have one extra GP contact costing £36 (Curtis and Burns 2016).

People who experienced common side effects were assumed to have one extra GP contact every 3 months costing £36 (Curtis and Burns 2016) and to consume a cost of £10 per year for medication relating to the management of common side effects (e.g. paracetamol or anti-inflammatory drugs for headaches).

14.2.11 Discounting

Costs and benefits were discounted at an annual rate of 3.5% in the second year of the Markov component of the model as recommended by (NICE 2014).

14.2.12 Handling uncertainty

Model input parameters were synthesised in a probabilistic analysis. This means that the input parameters were assigned probabilistic distributions (rather than being expressed as point estimates); this approach allowed more comprehensive consideration of the uncertainty characterising the input parameters and captured the non-linearity characterising the economic model structure. Subsequently, 10,000 iterations were performed, each drawing random values out of the distributions fitted onto the model input parameters. Results (mean costs and QALYs for each intervention) were averaged across the 10,000 iterations. This exercise provides more accurate estimates than those derived from a deterministic analysis (which utilises the mean value of each input parameter ignoring any uncertainty around the mean), by capturing the non-linearity characterising the economic model structure (Briggs, Sculpher et al. 2006).

The distributions of the odds ratios of relative effects of all treatments versus SSRIs or pill placebo (reflecting clinical management), as relevant, were obtained from the respective NMAs, defined directly from values recorded in each of the 10,000 iterations performed in WinBUGS.

Beta distribution was assigned to the following parameters: proportion of women in the study sample; the baseline risks of discontinuation and discontinuation due to side effects in those discontinuing; the proportion of people experiencing side effects; the probability of responders who did not remit moving to the remission state of the Markov model; and the probability of moving to specific relapse preventive treatments following successful completion of acute treatment. Utility values were also assigned a beta distribution after applying the method of moments on data reported in the relevant literature.

The 12-month probabilities of response and remission at various levels of symptom severity were given a beta distribution. The probabilities of response and remission following acute treatment, as well as the probability of remission and the baseline risk of relapse after a single (first) episode that were utilised in the Markov component of the model were determined by a Weibull distribution, as described earlier in methods. The probability distributions of the Weibull parameters (gamma and lambda) of recovery (‘baseline recovery’) that came from evidence synthesis in WinBUGS were defined directly from values recorded in each of 10,000 iterations performed in WinBUGS. This allowed the correlation between the Weibull parameters to be taken into account. The 12-month probabilities of

Update 2017

response and remission at various levels of symptom severity and the 12-month probability of 'baseline recovery' estimated from data synthesis were used to estimate hazard ratios of each parameter versus baseline recovery (see Table 293). These hazard ratios were then applied onto the 'baseline' lambda value obtained from data synthesis, in order to maintain the correlation between the lambda parameters for response and remission at each severity level and the gamma parameter that was estimated from data synthesis.

The hazard ratio of the risk of relapse for every additional depressive episode that was utilised in the Markov element of the model was given a log-normal distribution. The risk ratio of mortality was also assigned a log-normal distribution.

Uncertainty in intervention costs was taken into account by assigning probability distributions to the number of GP contacts and the number of individually delivered psychological therapy sessions. Different distributions around the number of GP contacts were used for people receiving active pharmacological interventions and for those receiving only clinical management (pill placebo). The number of therapist sessions per person attending group psychological interventions was not assigned a probability distribution because the number of group sessions remains the same, whether a participant attends the full course of treatment or a lower number of sessions. Drug acquisition costs were not given a probability distribution as these costs are set and characterised by minimal uncertainty. However, if people receiving maintenance pharmacological therapy attended fewer GP visits than the mode in the second year of maintenance treatment, then they were assumed to be prescribed smaller amounts of medication than optimal, and to subsequently incur lower drug acquisition costs. Unit costs of healthcare staff (GPs, clinical psychologists and PWPs) were assigned a normal distribution.

Healthcare costs associated with discontinuation of acute treatment and the states of relapse and remission in the Markov element of the model were assigned a gamma distribution.

Table 306 provides details on the types of distributions assigned to each input parameter and the methods employed to define their range.

A number of deterministic one-way sensitivity analyses were undertaken to explore the impact of alternative hypotheses on the results. The following scenarios were explored:

- Change in the number of previous episodes, resulting in a change in the risk of relapse in the Markov component of the model; the number of previous episodes was increased from 0 to 2 in people with less severe depression and was varied between 0 and 5 in people with more severe depression
- Use of higher utility values of 0.65 and 0.56 for less severe and more severe depression, respectively, reported in (Mann, Gilbody et al. 2009)
- Use of the values of 0.80 for remission and 0.62 for response not reaching remission reported in Koeser, Donisi et al. (2015)
- Setting the cost of GP visits associated with clinical management (pill placebo) at zero, in both the acute and maintenance phase of the model
- Changing the cost of relapse by ±50%
- Delivery of all psychological interventions by a band 5 PWP or a band 6 therapist (the unit cost of a band 6 therapist was estimated as the average of the unit costs of a band 5 PWP and a band 7 clinical psychologist)
- Delivery of group psychological interventions by band 7 clinical psychologists.
- Delivery of counselling in 8 sessions
- The effect of maintenance, relapse preventive treatment in people with more severe depression who remitted was zero and therefore all cohorts were subject to the (same) baseline risk of relapse.

Update 2017

- Change in baseline discontinuation (of citalopram in less severe depression and of sertraline in more severe depression) by ± 20%.

In addition, a probabilistic sensitivity analysis was run using data on response in completers derived from the bias-adjusted NMA models, which are described in Appendix N. Bias NMA models of the response in completers outcome in populations with both less and more severe depression suggested evidence of positive bias (i.e. overestimation of effect) in the comparisons of active versus inactive treatments in studies with larger variance (i.e. in smaller studies).

Table 306: Input parameters (deterministic values and probability distributions) that informed the economic models of interventions for the treatment of a new depressive episode in adults with less severe depression and adults with more severe depression

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
General characteristics of population			
Age of onset (years)	32	No distribution	Kessler, Berglund et al. (2005); Fernandez-Pujals, Adams et al. (2015)
Mean interval between episodes (years)	2	No distribution	GC advice
Number of previous episodes			GC expert opinion
- less severe depression	1	No distribution	GC expert advice
- more severe depression	3	No distribution	GP expert advice
Proportion of women	0.56	Beta: α=279; β=219	McManus, Bebbington et al. (2016); weighted prevalence of depression 2.9% in men, 3.7% in women, survey sample N=7,546
People with less severe depression: discontinuation - odds ratios vs citalopram			
Mirtazapine	0.542	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
BA	0.952		
Coping with Depression course (group)	1.755		
CBT individual (>15 sessions)	0.722		
CBT group (<15 sessions)	0.663		
IPT	0.860		
Short term PDPT individual	1.043		
Counselling	0.393		
Computerised CBT with support	1.275		
Computerised CBT without support	1.433		
Psychoeducational group programme	0.734		
Physical exercise programme	0.698		
CBT individual (>15 sessions) +citalopram	1.062		
IPT + citalopram	0.974		
Short-term PDPT + citalopram	0.783		
Physical exercise programme + sertraline	0.762		
Pill placebo	1.127		
People with less severe depression: discontinuation due to side effects in those discontinuing treatment – odds ratios vs SSRIs			

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
Mirtazapine	1.325	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
CBT + AD	3.506		
Short-term PDPT + AD	0.384		
Physical exercise programme + AD	0.427		
People with less severe depression: response in completers – odds ratios vs pill placebo			
Citalopram	2.188	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
Mirtazapine	3.255		
BA	3.761		
Coping with Depression course (group)	1.899		
CBT individual (>15 sessions)	3.992		
CBT group (<15 sessions)	2.686		
IPT	2.300		
Short term PDPT individual	2.439		
Counselling	1.860		
Computerised CBT with support	2.714		
Computerised CBT without support	1.597		
Psychoeducational group programme	1.451		
Physical exercise programme	2.992		
CBT individual (>15 sessions) +citalopram	1.963		
IPT + citalopram	9.047		
Short-term PDPT + citalopram	3.538		
Physical exercise programme + sertraline	2.015		
Wait list	0.373		
People with less severe depression: remission in completers – odds ratios vs pill placebo			
Citalopram	1.708	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
BA	3.894		
Coping with Depression course (group)	2.764		
CBT individual (>15 sessions)	2.662		
CBT group (<15 sessions)	2.015		
IPT	2.378		
Short term PDPT individual	1.614		

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
Counselling	3.882		
Computerised CBT with support	1.199		
Computerised CBT without support	1.411		
Psychoeducational group programme	1.723		
Physical exercise programme	0.906		
CBT individual (>15 sessions) +citalopram	1.182		
IPT + citalopram	3.565		
Short-term PDPT + citalopram	1.323		
Physical exercise programme + sertraline	0.988		
Wait list	0.211		
People with more severe depression: discontinuation - odds ratios vs sertraline			
Mirtazapine	0.820	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
BA	1.959		
CBT individual (>15 sessions)	0.693		
CBT group (<15 sessions)	0.492		
Short term PDPT individual	0.361		
Counselling	0.269		
CBT individual (>15 sessions) + sertraline	0.215		
Pill placebo	1.148		
People with more severe depression: discontinuation due to side effects in those discontinuing treatment – odds ratios vs SSRIs			
Mirtazapine	1.438	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
CBT + AD	0.185		
People with more severe depression: response in completers – odds ratios vs pill placebo			
Sertraline	1.839	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
Mirtazapine	2.811		
BA	27.363		
CBT individual (>15 sessions)	1.553		
CBT group (<15 sessions)	3.106		
Short term PDPT individual	5.507		
Counselling	4.821		

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
CBT individual (>15 sessions) + sertraline Wait list	6.940 0.651		
People with more severe depression: remission in completers – odds ratios vs pill placebo			
Sertraline Mirtazapine CBT individual (>15 sessions) Short term PDPT individual CBT individual (>15 sessions) + sertraline Wait list	1.425 0.882 1.718 1.469 6.811 0.055	Based on NMA	Guideline NMA; distribution based on 10,000 iterations
Baseline risk of discontinuation			
Less severe depression – citalopram More severe depression – sertraline	0.370 0.340	Beta: α=185; β=315 Beta: α=170; β=330	Based on a review of studies (Bull, Hunkeler et al. 2002, Hansen, Vach et al. 2004, Lewis, Marcus et al. 2004, Olfson, Marcus et al. 2006, Goethe, Woolley et al. 2007, Burton, Anderson et al. 2012) and further expert opinion
Baseline risk of discontinuation due to side effects in those discontinuing			
Less severe depression – SSRIs More severe depression – SSRIs	0.405 0.441	Beta: α=203; β=297 Beta: α=221; β=279	Based on data on discontinuation due to side effects reported in Goethe, Woolley et al. (2007) and Bull, Hunkeler et al. (2002) for SSRIs, using the estimated baseline risk of discontinuation of SSRIs for less and more severe depression and assuming that discontinuation due to side effects is independent of depressive symptom severity
Response and remission in completers – pill placebo			
Less severe depression – response Less severe depression – remission More severe depression – response More severe depression – remission Hazards ratios of the above states versus 12-month baseline probability of recovery were estimated using the probabilities below: 12-month response	0.505 0.491 0.492 0.341 0.793	Based on Weibull parameters (lambda and gamma) for baseline probability of recovery [shown below] Beta: α=235; β=61	Synthesis of data from Holma, Holma et al. (2008) Keller and Shapiro (1981) Keller, Klerman et al. (1984) Keller, Lavori et al. (1992) Mueller, Keller et al. (1996) Skodol, Grilo et al. (2011) Stegenga, Kamphuis et al. (2012) Gonzales, Lewinsohn et al. (1985)

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
– mild depression	0.677	Beta: α=265; β=126	using a Bayesian approach – random effects model
– moderate depression	0.725	Beta: α=233; β=88	Simon, Goldberg et al. (1999). For less severe depression the mean values of mild and moderate depression were used.
– severe depression			
12-month remission			
– mild depression	0.793	Beta: α=235; β=61	
– moderate depression	0.645	Beta: α=252; β=139	
– severe depression	0.549	Beta: α=176; β=145	
Probability of responders (without remission) moving to remission Markov state			
– less severe depression	0.60	Beta: α=60; β=40	Based on GC expert opinion
– more severe depression	0.30	Beta: α=30; β=70	
Probability of developing common side effects			
– SSRIs alone or in combination	0.12	Beta: α=2,752; β=20,868	Anderson, Pace et al. (2012)
– mirtazapine	0.16	Beta: α=147; β=754	
Probability of moving to specific relapse preventive treatment according to acute treatment received – more severe depression			
Acute drug -> maintenance drug	0.80	Beta: α=80; β=20	Based on GC expert opinion
Acute psych -> maintenance 4 sessions	0.50	Beta: α=50; β=50	
Acute combined -> maintenance drug	0.80	Beta: α=80; β=20	
Baseline risk of relapse after a single (first) episode			Synthesis of data from Eaton, Shao et al. (2008) and Mattisson, Bogren et al. (2007), using a Bayesian approach – fixed effects model
Weibull distribution – lambda	0.095	WinBUGS output	
Weibull distribution – gamma	0.611	WinBUGS output	
Hazard ratio – new vs previous episode	1.15	Log-normal: 95% CI 1.11 to 1.18	Kessing and Andersen (1999)
Baseline probability of recovery			Synthesis of data from
Weibull distribution – lambda	1.171	WinBUGS output	Holma, Holma et al. (2008)
Weibull distribution – gamma	0.440	WinBUGS output	Keller and Shapiro (1981)
			Keller, Klerman et al. (1984)
			Keller, Lavori et al. (1992)
			Mueller, Keller et al. (1996)

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
			Skodol, Grilo et al. (2011) Stegenga, Kamphuis et al. (2012) Gonzales, Lewinsohn et al. (1985) using a Bayesian approach – random effects model
Mortality Risk ratio – depressed vs non-depressed Baseline mortality – non-depressed	 1.52 Age/sex specific	 Log-normal: 95% CI 1.45 to 1.59 No distribution	Cuijpers, Vogelzangs et al. (2014) General mortality statistics for the UK population (ONS 2015)
Utility values Less severe depression More severe depression Remission Response not reaching remission Disutility due to side effects Remission state in Markov component	 0.60 0.42 0.85 0.72 0.09 0.81	 Beta: α=182; β=122 Beta: α=54; β=75 Beta: α=923; β=163 Beta: α=123; β=48 Beta: α=6; β=59 Beta: α=531; β=125	Distributions determined using method of moments, based on data reported in (Sapin, Fantino et al. 2004, Sullivan, Valuck et al. 2004, Sobocki, Ekman et al. 2006, Sobocki, Ekman et al. 2007) and further assumptions
Intervention costs – resource use COMPLETERS Number of GP contacts – drug treatment – Acute treatment – Discontinuation due to side effects – Side effects during acute treatment – Side effects during maintenance treatment – 1st year maintenance – 2nd year maintenance – tapering Number of GP contacts – clinical management – Acute treatment – 1st year maintenance	 4 1 1 4 6 3 1 4 3	 0.70: 4, 0.30: 2-3 0.80: 1, 0.20: 0 No distribution assigned 2 or 4 in second year 0.70: 6, 0.20: 4-5, 0.10: 2-3 0.70: 3, 0.30: 1-2 0.70: 1, 0.30: 2 0.50: 4, 0.50: 2-3 0.70: 3, 0.20: 1-2, 0.10: 0	Probabilities assigned to numbers of sessions Number of visits based on GC expert opinion; probabilities based on assumption. If number of GP visits in 2nd year of maintenance pharmacological treatment was lower than 3, only 50% of the drug acquisition cost was incurred and 50% of annual GP contacts due to side effects were made See note on GP visits in 2nd year of maintenance drug treatment

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
– 2nd year maintenance	1	0.70: 1, 0.30: 0	
Number of GP contacts – psych therapy			Details on costs of psychological therapies are provided in Table 301 and Table 303.
– Acute treatment	1	No distribution	
– Maintenance treatment	2	0.60: 2, 0.40: 1	
Acute psychological therapies – number of sessions			
cCBT with support	5	0.60: 5, 0.20: 4, 0.20: 2-3	cCBT with/without support: fixed digital therapy provider + capital cost of £49.2 added to the therapist cost. For cCBT with support one extra initial (longer) visit added to the 5 visits.
cCBT without support	0	No distribution	
Coping with Depression group	12	No distribution	Participants missing one or more group sessions assumed not to be replaced by others; therefore no impact on total intervention cost
Psychoeducational group	9	No distribution	
CBT group	9	No distribution	
CBT individual	16	0.60: 16, 0.40: 5-15	Number of visits based on GC expert opinion; probabilities based on assumption
BA	16	0.60: 16, 0.40: 5-15	
IPT	16	0.60: 16, 0.40: 5-15	
Short-term PDPT	16	0.60: 16, 0.40: 5-15	
Counselling	16	0.60: 16, 0.40: 5-15	
Maintenance psychological therapies – number of sessions			
MBCT (group)	12	No distribution	Number of visits based on GC expert opinion; probabilities based on assumption
CBT group	4	No distribution	
4 individual sessions	4	0.60: 4, 0.40: 2-3	
Physical exercise programme	15	No distribution	
DISCONTINUERS (acute treatment)			
Number of GP contacts – drug treatment or clinical management	1	No distribution	One pack of drugs assumed to be consumed by those discontinuing acute drug treatment
Number of GP contacts – psych therapy			

Update 2017

Input parameter	Mean deterministic value	Probability distribution	Source of data - comments
Number of psychological therapy sessions	1	No distribution	
– cCBT with support	1	No distribution	Plus initial visit and full fixed cost of programme provider
– cCBT without support	0	No distribution	Plus full fixed cost of programme provider
– Coping with Depression group	12	No distribution	
– Psychoeducational group	9	No distribution	People discontinuing group therapy or physical exercise programme were assumed to incur the full cost of therapy
– CBT group	9	No distribution	
– CBT individual	4	No distribution	
– BA	4	No distribution	
– IPT	4	No distribution	
– Short-term PDPT	4	No distribution	
– Counselling	4	No distribution	
Number of sessions – exercise programme	15	No distribution	
Intervention costs - unit costs	See Table 298	No distribution	National drug tariff, January 2017 (NHS 2017)
Drug acquisition costs	£36	Normal, SE=0.05*mean	Curtis and Burns (2016); distribution based on assumption
GP unit cost	£97	Normal, SE=0.05*mean	See Table 299; distribution based on assumption
Clinical psychologist unit cost	£42	Normal, SE=0.05*mean	See Table 300; distribution based on assumption
PWP unit cost			
Annual NHS health state cost	£1,483	Gamma	Based primarily on cost data reported in Byford, Barrett et al. (2011), supplemented by data from (Radhakrishnan, Hammond et al. 2013, Curtis and Burns 2016, NHS-England 2016), expressed in 2016 prices using the HCHS inflation index (Curtis and Burns 2016). Distribution based on assumption
Relapse - remaining in state	£1,079	SE=0.20*mean	
Relapse - final year before remission	£493		
Remission	£141		
Remission – 1st year extra cost	£228		
Cost of treatment after discontinuation			
Annual discount rate	0.035	No distribution	Applied to both costs and outcomes. (NICE 2014)

Update 2017

1

14.2.131 Presentation of the results

2 Results of the economic analysis are presented as follows:

3 Results are reported separately for each cohort examined in the economic model. In each
4 analysis, mean total costs and QALYs are presented for each intervention, averaged across
5 10,000 iterations of the model. An incremental analysis is provided for each cohort, in table
6 format, where all options have been listed from the most to the least effective (in terms of
7 QALYs gained). Options that are dominated by absolute dominance (that is, they are less
8 effective and more costly than one or more other options) or by extended dominance (that is,
9 they are less effective and more costly than a linear combination of two alternative options)
10 are excluded from further analysis. Subsequently, incremental cost-effectiveness ratios
11 (ICERs) are calculated for all pairs of consecutive options remaining in analysis.

12 ICERs are calculated by the following formula:

13 $$ICER = \Delta C / \Delta E$$

14 where ΔC is the difference in total costs between two interventions and ΔE the difference in
15 their effectiveness (QALYs). ICERs express the extra cost per extra unit of benefit (QALY)
16 associated with one treatment option relative to its comparator. The treatment option with the
17 highest ICER below the NICE lower cost effectiveness threshold of £20,000/QALY (NICE
18 2008) is the most cost-effective option.

19 In addition to ICERs, the mean net monetary benefit (NMB) of each intervention is presented.
20 This is defined by the following formula:

21 $$NMB = E \bullet \lambda - C$$

22 where E and C are the effectiveness (number of QALYs) and costs associated with the
23 treatment option, respectively, and λ is the level of the willingness-to-pay (WTP) per unit of
24 effectiveness, set at the NICE lower cost effectiveness threshold of £20,000/QALY (NICE
25 2008). The intervention with the highest NMB is the most cost-effective option (Fenwick,
26 Claxton et al. 2001).

27 Incremental mean costs and effects (QALYs) of each intervention versus clinical
28 management (pill placebo) are also presented in the form of cost effectiveness planes.

29 The probability of each intervention being the most cost-effective option at the NICE lower
30 cost effectiveness threshold of £20,000/QALY is also provided, calculated as the proportion
31 of iterations (out of the 10,000 iterations run) in which the intervention had had the highest
32 NMB among all interventions considered in the analysis.

33 The probability of each intervention being the most cost-effective option at the NICE lower
34 cost effectiveness threshold of £20,000/QALY is also provided in a step-wise approach,
35 according to which the most cost-effective intervention is omitted at each step and the
36 probability of the intervention with the next highest NMB is re-calculated.

37 The mean ranking in terms of cost effectiveness is also reported for each intervention (out of
38 the 10,000 iterations run), with lower rankings suggesting higher cost effectiveness. Mean
39 rankings are also provided in a step-wise approach.

40 ICERs (or cases of dominance) are also provided for every treatment option versus the next
41 most cost-effective one.

42 The probabilities of each intervention being cost-effective at various cost effectiveness
43 thresholds are illustrated in cost-effectiveness acceptability curves (CEACs). Finally, the
44 cost-effectiveness acceptability frontiers (CEAFs) were also plotted; these show the

Update 2017

treatment option with the highest mean NMB over different cost effectiveness thresholds, and the probability that the option with the highest NMB is the most cost-effective among those assessed (Fenwick, Claxton et al. 2001).

14.2.14 Validation of the economic model

The economic model (including the conceptual model and the identification and selection of input parameters) was developed by the health economist in collaboration with a health economics sub-group formed by members of the Guideline Committee. As part of the model validation, all inputs and model formulae were systematically checked; 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 base-case results and results of sensitivity analyses were discussed with the Guideline Committee to confirm their plausibility. In addition, the economic model (excel spreadsheet) and this chapter were checked for their validity and accuracy by a health economist that was external to the guideline development team.

14.3 Economic modelling results

14.3.1 Adults with less severe depression

The base-case results of the economic analysis are provided in Table 307. This table provides mean QALYs and mean intervention and total costs for each intervention assessed in the economic analysis, as well as the results of incremental analysis, the mean NMB of each intervention, and its ranking by cost effectiveness (with higher NMBs and lower rankings indicating higher cost effectiveness). Interventions have been ordered from the most to the least effective in terms of number of QALYs gained. Intervention costs include costs for treatment completers and costs for those who discontinued treatment. According to the results, CBT individual was the most effective intervention in terms of QALYs gained, followed by IPT combined with citalopram, and behavioural activation. Mirtazapine and CBT group were also included in the top five effective interventions. Clinical management, reflecting pill placebo trial arms, was the least effective intervention. In terms of cost-effectiveness, mirtazapine appears to be the best treatment option (highest mean NMB), followed by CBT group, physical exercise programme, citalopram, and cCBT with support. Other low-intensity interventions, such as physical exercise programme combined with sertraline, psychoeducational group programme, coping with depression group course and cCBT without or with minimal support also ranked highly. These were followed by high intensity psychological interventions alone or in combination, and by clinical management, in the following order: CBT individual, behavioural activation, IPT combined with citalopram, clinical management, IPT, short term PDPT, short term PDPT combined with citalopram, counselling, and CBT individual combined with citalopram. The probability of mirtazapine being the most cost-effective option was 0.45 at the NICE lower cost effectiveness threshold of £20,000/QALY.

Table 307: Results of economic modelling: interventions for people with a new episode of less severe depression – base-case analysis (mean values from probabilistic analysis)

Acute treatment option	Mean per person			ICER (£/QALY)	NMB / person	Prob best[1]	Mean rank
	QALY	Interv cost	Total cost				
CBT individual	1.682	£1,058	£2,563	£102,810	£31,079	0.003	10.20
IPT + citalopram	1.682	£1,052	£2,585	dominated	£31,052	0.022	10.52
BA	1.680	£1,007	£2,531	ext domin	£31,064	0.005	10.42

Update 2017

Acute treatment option	Mean per person			ICER (£/QALY)	NMB / person	Prob best[1]	Mean rank
	QALY	Interv cost	Total cost				
Mirtazapine	1.673	£107	£1,658		£31,808	0.448	3.89
CBT group	1.672	£130	£1,674	dominated	£31,758	0.270	2.93
IPT	1.668	£1,028	£2,593	dominated	£30,775	0.000	13.63
Short-term PDPT +cital	1.668	£1,098	£2,679	dominated	£30,680	0.005	13.98
Counselling	1.667	£1,130	£2,685	dominated	£30,660	0.009	13.61
Physical exercise prog	1.664	£94	£1,669	dominated	£31,616	0.079	3.96
Short term PDPT	1.664	£986	£2,576	dominated	£30,697	0.000	14.35
Citalopram	1.661	£96	£1,710	dominated	£31,511	0.030	5.17
cCBT with support	1.657	£156	£1,781	dominated	£31,354	0.022	6.84
Coping with Depression	1.655	£202	£1,847	dominated	£31,245	0.008	8.22
Exercise + sertraline	1.655	£159	£1,790	dominated	£31,301	0.017	7.67
Psychoeducation	1.652	£130	£1,757	dominated	£31,287	0.071	7.79
CBT individual + cital	1.649	£1,014	£2,678	dominated	£30,303	0.000	16.67
cCBT	1.649	£85	£1,746	dominated	£31,234	0.011	8.37
Clinical management	1.632	£77	£1,760	dominated	£30,871	0.000	12.79

Notes:

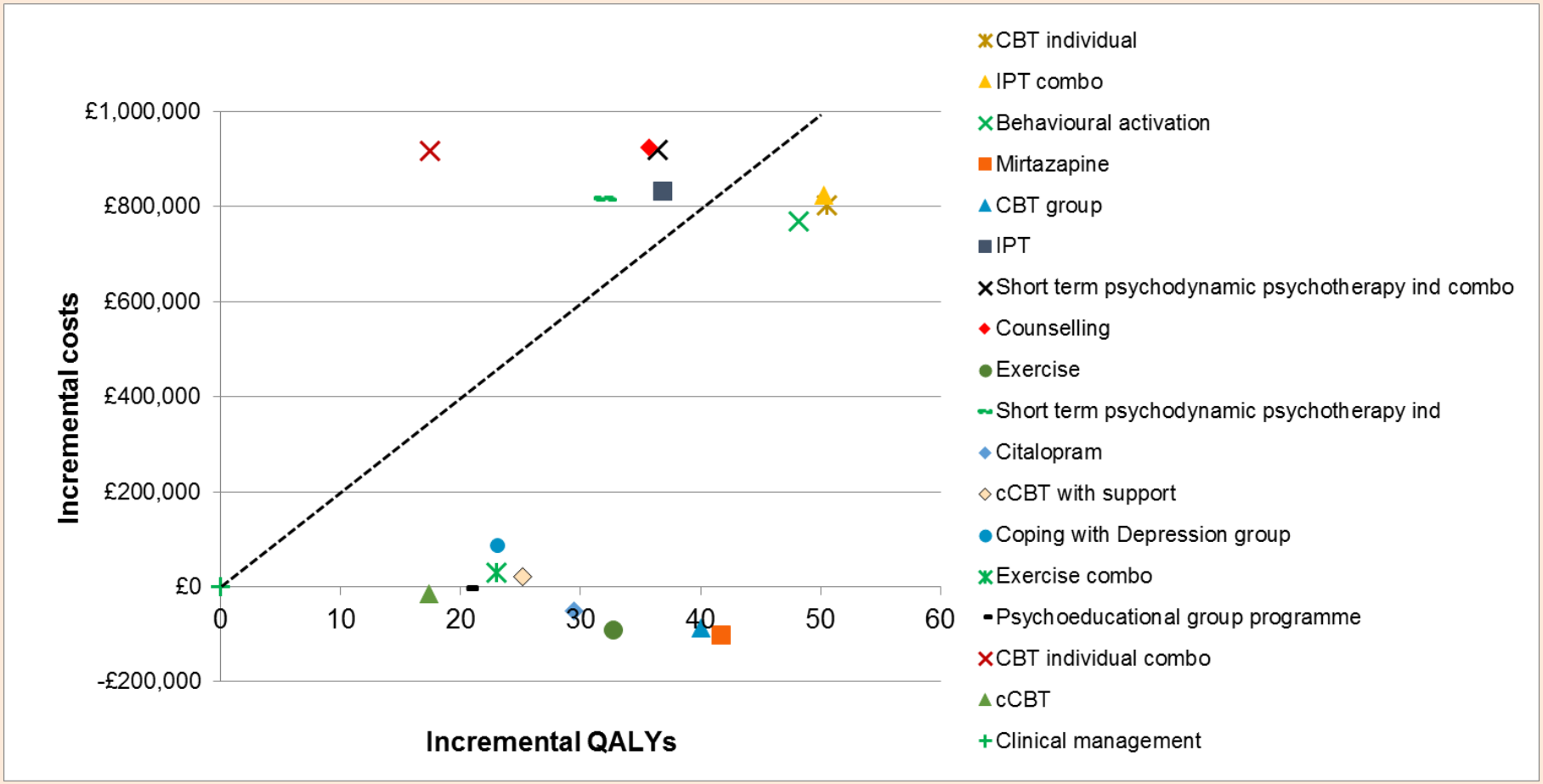
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY

BA: behavioural activation; CBT: cognitive behavioural therapy; cCBT: computerised cognitive behavioural therapy; cital: citalopram; ext domin: extendedly dominated; ICER: incremental cost effectiveness ratio; interv: intervention; IPT: interpersonal psychotherapy; NMB: net monetary benefit; PDPT: psychodynamic psychotherapy; Prob: probability; prog: programme

Update 2017

Figure 39 provides the cost effectiveness plane of the analysis. Each intervention is placed on the plane according to its incremental costs and QALYs compared with clinical management (pill placebo), which is placed at the origin. The slope of the dotted line indicates the NICE lower cost effectiveness threshold, suggesting that IPT, short term PDPT, counselling, short term PDPT combined with citalopram and CBT individual combined with citalopram are not cost-effective compared with clinical management (since they all lie on the left side of the dotted line).

Figure 39. Cost effectiveness plane of interventions for the treatment of a new episode of less severe depression in adults plotted against clinical management (pill placebo) – incremental costs and QALYs versus clinical management per 1,000 adults with less severe depression



Update 2017

Table 308 presents the interventions ordered from the most to the least cost-effective at the NICE lower cost effectiveness threshold (£20,000/QALY), the incremental cost effectiveness between each option and the next most cost-effective option (in terms of the ICER of the most effective intervention versus its comparator or cases of dominance), and the probabilities and mean rankings of cost effectiveness among all available treatment options obtained in a step-wise approach, after the most cost-effective intervention is omitted from analysis and the probability and mean ranking of the next most cost-effective option among the remaining available treatment options are re-calculated. It can be seen that, with the exception of mirtazapine and CBT group, the next most cost-effective interventions up to (and including) short term PDPT have probabilities of being cost-effective among remaining options that are lower than 0.40, although increasingly fewer interventions are included in the analysis, indicating the uncertainty characterising the results.

Table 308: Results of economic modelling: interventions for adults with a new episode of less severe depression – probability of being best and mean ranking at the NICE lower cost effectiveness threshold (step-wise approach)

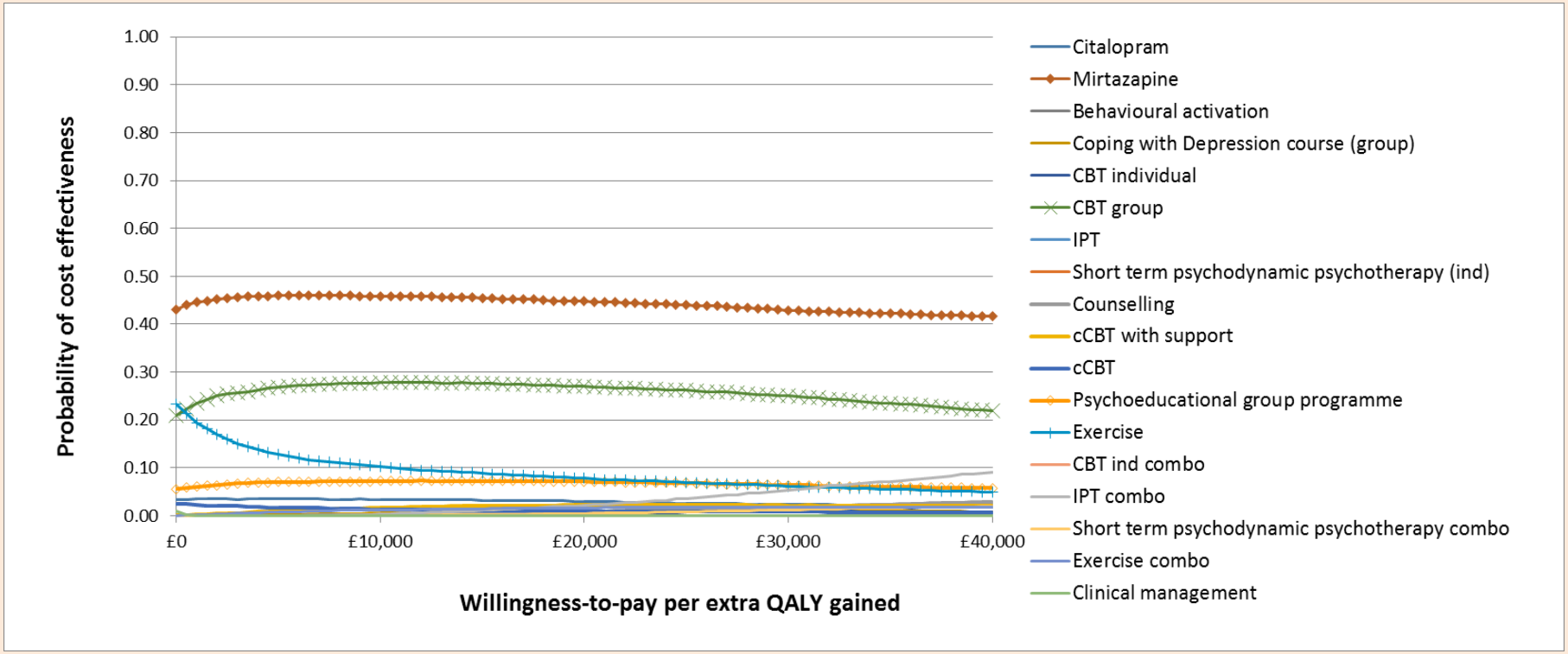
Acute treatment option	Incremental cost effectiveness (each option vs next most cost-effective option)	Probability being best[1] (step-wise approach)	Mean ranking (step-wise approach)
Mirtazapine	Mirtazapine dominant	0.448	3.89
CBT group	£737/QALY	0.465	2.40
Exercise	Exercise dominant	0.312	2.64
Citalopram	Citalopram dominant	0.276	3.04
cCBT with support	cCBT with support dominant	0.172	3.80
Exercise + sertraline	£14,020/QALY	0.197	3.84
Psychoeducation	Coping with Depression vs psychoeducation £36,963/QALY	0.317	3.80
Coping with Depression group	£17,974/QALY	0.221	3.11
cCBT	CBT ind vs cCBT £24,656/QALY	0.375	2.81
CBT individual	£13,709/QALY	0.221	3.10
BA	IPT + cital vs BA £25,820/QALY	0.285	2.71
IPT + citalopram	£16,400/QALY	0.388	2.44
Clinical management	IPT vs clinical management £22,612/QALY	0.268	2.40
IPT	£3,587/QALY	0.287	2.45
Short term PDPT	Short term PDPT + cital vs short term PDPT £23,890/QALY	0.232	2.17
Short term PDPT + citalopram	Short term PDPT + citalopram dominant	0.429	1.76
Counselling	£413/QALY	0.711	1.29
CBT individual + citalopram		1.000	1.00

Notes:
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
BA: behavioural activation; CBT: cognitive behavioural therapy; cCBT: computerised cognitive behavioural therapy; cital: citalopram; IPT: interpersonal psychotherapy; PDPT: psychodynamic psychotherapy

The CEAC and CEAF of the analysis are shown in Figure 40 and Figure 41 respectively. It can be seen that mirtazapine is the most cost-effective option at any cost effectiveness threshold between zero and £40,000/QALY, with a probability that ranges between 0.42 and 0.46.

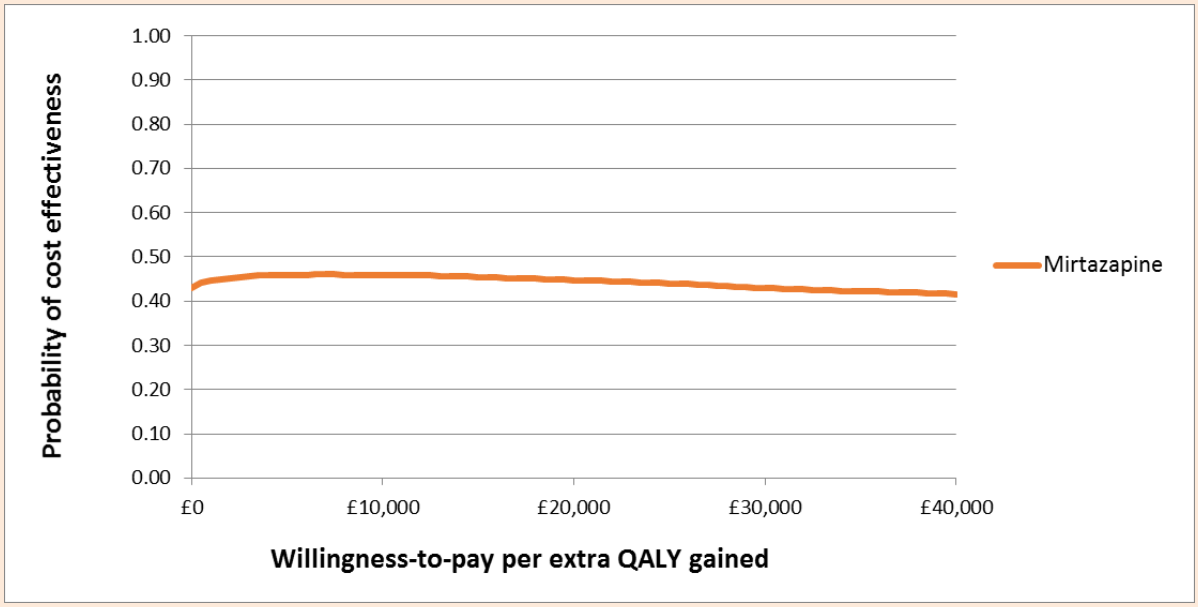
Update 2017

Figure 40. Cost-effectiveness acceptability curves of interventions for the treatment of a new episode of less severe depression in adults



Update 2017

Figure 41 Cost-effectiveness acceptability frontier of interventions for the treatment of a new episode of less severe depression in adults



Results were robust to alternative scenarios tested in one-way deterministic sensitivity analysis, with the following exceptions:

- When the higher utility value from Mann, Gilbody et al. (2009) was attached to less severe depression (translating into a more limited scope for HRQoL improvement following successful treatment), the ranking of the 6 highest cost-effective interventions did not change; however, cCBT without or with minimal support became more cost-effective than combined exercise + sertraline and Coping with Depression course (group). Moreover, IPT combined with citalopram, BA and CBT individual became less cost-effective than clinical management.
- When the cost of relapse was assumed to be 50% lower than the base-case value, the raking of the 6 highest cost-effective interventions did not change; however, cCBT without or with minimal support became more cost-effective than combined exercise + sertraline and Coping with Depression course (group). Moreover, IPT combined with citalopram, BA and CBT individual became less cost-effective than clinical management.
- When all psychological interventions were assumed to be delivered by a band 5 PWP, the intervention cost of individual high-intensity psychological interventions was reduced and their relative cost effectiveness increased, resulting in changes in ranking. According to this scenario, the order of interventions from the most to the least cost-effective in deterministic analysis was as follows: mirtazapine, CBT group, physical exercise programme, CBT individual, IPT combined with citalopram, BA, citalopram, psychoeducational group programme, cCBT with support, cCBT without or with minimal support, physical exercise programme combined with sertraline, coping with Depression course (group), counselling, IPT, short term PDPT combined with citalopram, short term PDPT, clinical management, CBT individual combined with citalopram. Assuming that individual high-intensity psychological interventions were delivered by a band 6 therapist had a less profound impact on the results, but still improved the cost effectiveness of individual high-intensity psychological interventions, all of which became more cost-effective than pill placebo. It needs to be noted that combining a scenario of delivery of individual psychological interventions by a band 5 PWP with delivery of group psychological interventions by a band 7 clinical psychologist had no impact on the cost effectiveness of group CBT, which remained the second most cost-effective option.

Update 2017

- When counselling was assumed to be delivered in 8 sessions instead of 16, it became the 10th most cost-effective option, following cCBT without or with minimal support.
- When the baseline treatment discontinuation for citalopram was changed by ±20%, there were small changes in the ranking of interventions, although the order of the first 6 most cost-effective interventions remained the same.

The results of the probabilistic sensitivity analysis that utilised data on response in completers from the respective bias NMA model are shown in Table 309. It can be seen that results only modestly changed: IPT combined with citalopram was the most effective intervention under this analysis, followed by CBT individual, BA, short-term PDPT combined with citalopram and IPT. Clinical management remained the least effective intervention. Regarding cost effectiveness, mirtazapine was again the most cost-effective intervention (with a 0.37 probability of being cost-effective at the NICE lower cost-effectiveness threshold of £20,000/QALY), followed by physical exercise programme, citalopram, CBT group, psychoeducational group programme, cCBT with support, coping with Depression course group, physical exercise programme combined with sertraline and cCBT without or with minimal support. These low-cost interventions, which were also the most cost-effective interventions in the base-case analysis, were followed by high-intensity psychological interventions and clinical management in the following order: IPT combined with citalopram, clinical management, CBT individual, BA, IPT, short term PDPT combined with citalopram, short term PDPT, counselling, and, finally, CBT individual combined with citalopram, which was also the least cost-effective intervention in base-case analysis. It is noted that CBT individual and BA appear to be less cost-effective than clinical management in this sensitivity analysis, whilst they were more cost-effective than clinical management in the base-case analysis.

Table 309: Results of economic modelling: interventions for people with a new episode of less severe depression – sensitivity analysis based on bias-adjusted NMA models (mean values from probabilistic analysis)

Acute treatment option	Mean per person			ICER (£/QALY)	NMB / person	Prob best[1]	Mean rank
	QALY	Interv cost	Total cost				
IPT + citalopram	1.681	£1,048	£2,582	£40,624	£31,039	0.043	9.35
CBT individual	1.676	£1,056	£2,583	dominated	£30,935	0.005	10.53
BA	1.672	£1,005	£2,557	ext domin	£30,877	0.005	11.28
Short-term PDPT +cital	1.662	£1,098	£2,698	dominated	£30,546	0.007	14.03
IPT	1.661	£1,023	£2,617	dominated	£30,597	0.000	14.19
Mirtazapine	1.660	£107	£1,708	£6,556	£31,483	0.366	5.41
Exercise	1.657	£94	£1,694		£31,454	0.099	4.15
Short term PDPT	1.657	£985	£2,600	dominated	£30,540	0.000	14.63
CBT group	1.657	£129	£1,733	dominated	£31,398	0.122	4.96
Citalopram	1.657	£96	£1,724	dominated	£31,407	0.069	4.72
cCBT with support	1.655	£156	£1,782	dominated	£31,326	0.075	5.64
Counselling	1.655	£1,129	£2,731	dominated	£30,366	0.008	14.72
Psychoeducation	1.654	£129	£1,747	dominated	£31,328	0.146	6.00
Coping with Depression	1.652	£202	£1,855	dominated	£31,183	0.018	7.54
Exercise + sertraline	1.648	£159	£1,813	dominated	£31,144	0.018	8.12
cCBT	1.645	£85	£1,760	dominated	£31,136	0.019	8.23
CBT individual + cital	1.643	£1,012	£2,698	dominated	£30,154	0.000	16.65
Clinical management	1.634	£77	£1,746	dominated	£30,936	0.000	10.82
Notes:							

Update 2017

Acute treatment option	Mean per person			ICER (£/QALY)	NMB / person	Prob best[1]	Mean rank
	QALY	Interv cost	Total cost				
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY BA: behavioural activation; CBT: cognitive behavioural therapy; cCBT: computerised cognitive behavioural therapy; cital: citalopram; ext domin: extendedly dominated; ICER: incremental cost effectiveness ratio; interv: intervention; IPT: interpersonal psychotherapy; NMB: net monetary benefit; PDPT: psychodynamic psychotherapy; Prob: probability; prog: programme							

14.3.21 Adults with more severe depression

The base-case results of the economic analysis are provided in Table 310. This table provides mean QALYs and mean intervention and total costs for each intervention assessed in the economic analysis, as well as the results of incremental analysis, the mean NMB of each intervention, and its ranking by cost effectiveness (with higher NMBs and lower rankings indicating higher cost effectiveness). Interventions have been ordered from the most to the least effective in terms of number of QALYs gained. Intervention costs include costs for treatment completers and costs for those who discontinued treatment. According to the results, CBT individual combined with sertraline was the most effective intervention in terms of QALYs gained, followed by BA, short term PDPT, CBT group and counselling. Clinical management, reflecting pill placebo trial arms, was the least effective intervention with the exception of cCBT without or with minimal support, which was the only intervention ranked below clinical management in terms of effectiveness. CBT individual combined with sertraline was also the most cost-effective intervention among those assessed, followed by CBT group, BA, sertraline, physical exercise programme, short term PDPT, mirtazapine, counselling, CBT individual and, clinical management, and cCBT without or with minimal support. The probability of CBT individual combined with sertraline being the most cost-effective option was 0.31 at the NICE lower cost effectiveness threshold of £20,000/QALY.

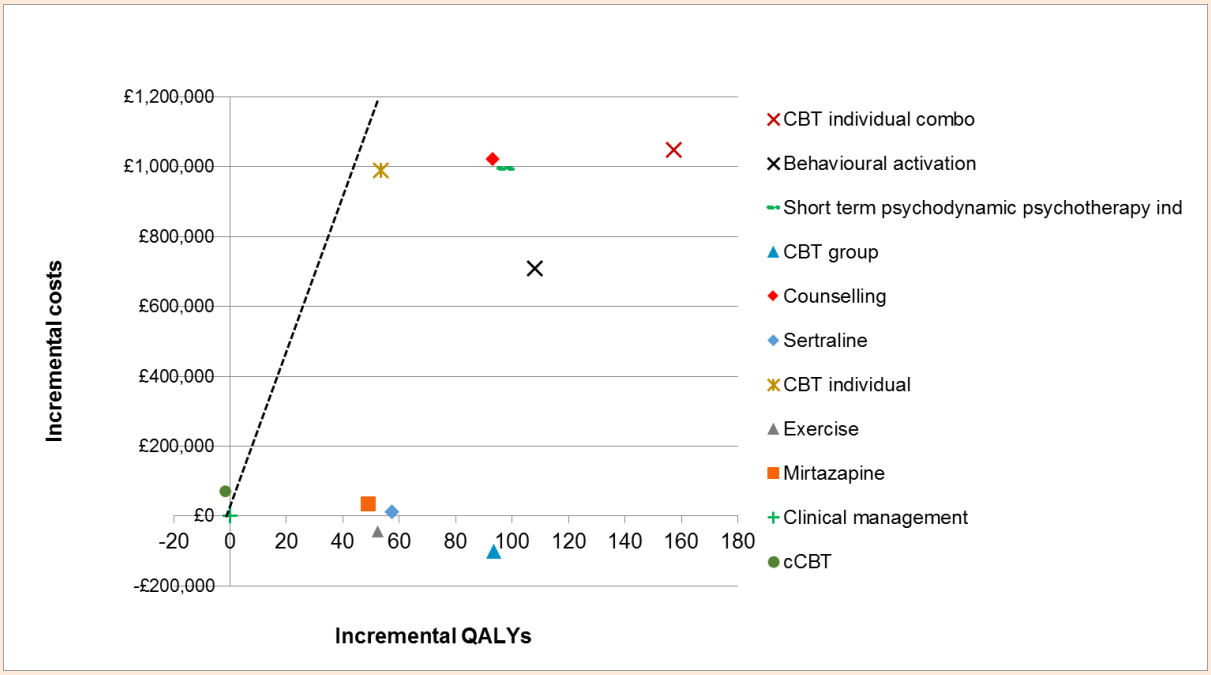
Table 310: Results of economic modelling: interventions for people with a new episode of more severe depression – base-case analysis (mean values from probabilistic analysis)

Acute treatment option	Mean per person			ICER (£/QALY)	NMB / person	Prob best[1]	Rank
	QALY	Interv cost	Total cost				
CBT individual + sertral	1.536	£1,313	£3,063	£18,026	£27,660	0.311	3.11
BA	1.487	£884	£2,724	ext domin	£27,016	0.130	4.86
Short term PDPT	1.477	£1,172	£3,009	Dominated	£26,525	0.034	5.98
CBT group	1.472	£130	£1,914		£27,534	0.238	3.23
Counselling	1.472	£1,195	£3,038	dominated	£26,402	0.047	6.36
Sertraline	1.436	£99	£2,028	dominated	£26,695	0.050	4.93
CBT individual	1.432	£1,086	£3,005	dominated	£25,639	0.000	8.37
Physical exercise prog	1.431	£94	£1,971	dominated	£26,655	0.133	5.72
Mirtazapine	1.428	£104	£2,049	dominated	£26,505	0.056	5.72
Clinical management	1.379	£80	£2,015	Dominated	£25,562	0.000	8.68
cCBT	1.377	£85	£2,085	Dominated	£25,457	0.001	9.05
Notes: 1 at the NICE lower cost-effectiveness threshold of £20,000/QALY BA: behavioural activation; CBT: cognitive behavioural therapy; cCBT: computerised cognitive behavioural therapy; ext domin: extendedly dominated; ICER: incremental cost effectiveness ratio; interv: intervention; NMB: net monetary benefit; PDPT: psychodynamic psychotherapy; Prob: probability; prog: programme; sertral: sertraline							

Update 2017

Figure 42 provides the cost-effectiveness plane of the analysis. Each intervention is placed on the plane according to its incremental costs and QALYs compared with clinical management (pill placebo), which is placed at the origin. The slope of the dotted line indicates the NICE lower cost effectiveness threshold, suggesting that all interventions assessed are cost-effective compared with clinical management, with the exception of cCBT without or with minimal support (since this is the only intervention that lies on the left side of the dotted line).

Figure 42. Cost-effectiveness plane of interventions for the treatment of a new episode of more severe depression in adults plotted against clinical management (pill placebo) – incremental costs and QALYs versus clinical management per 1,000 adults with more severe depression



Update 2017

Table 311 presents the interventions ordered from the most to the least cost-effective at the NICE lower cost effectiveness threshold (£20,000/QALY), the incremental cost effectiveness between each option and the next most cost-effective option (in terms of the ICER of the most effective intervention versus its comparator or cases of dominance), and the probabilities and mean rankings of cost effectiveness among all available treatment options obtained in a step-wise approach, after the most cost-effective intervention is omitted from analysis and the probability and mean ranking of the next most cost-effective option among the remaining available treatment options are re-calculated. It can be seen that all interventions up to short-term psychodynamic psychotherapy have probabilities of being cost-effective among remaining options that are lower than 0.40, although increasingly fewer interventions are included in the analysis, indicating the uncertainty characterising the results in the middle ranking cost effectiveness area.

Table 311: Results of economic modelling: interventions for adults with a new episode of more severe depression – probability of being best and mean ranking at the NICE lower cost effectiveness threshold (step-wise approach)

Acute treatment option	Incremental cost effectiveness (each option vs next most cost-effective option)	Probability being best[1] (step-wise approach)	Mean ranking (step-wise approach)
CBT individual + sertr	£18,026/QALY	0.311	3.11
CBT group	BA vs CBT group £55,434/QALY	0.368	2.70

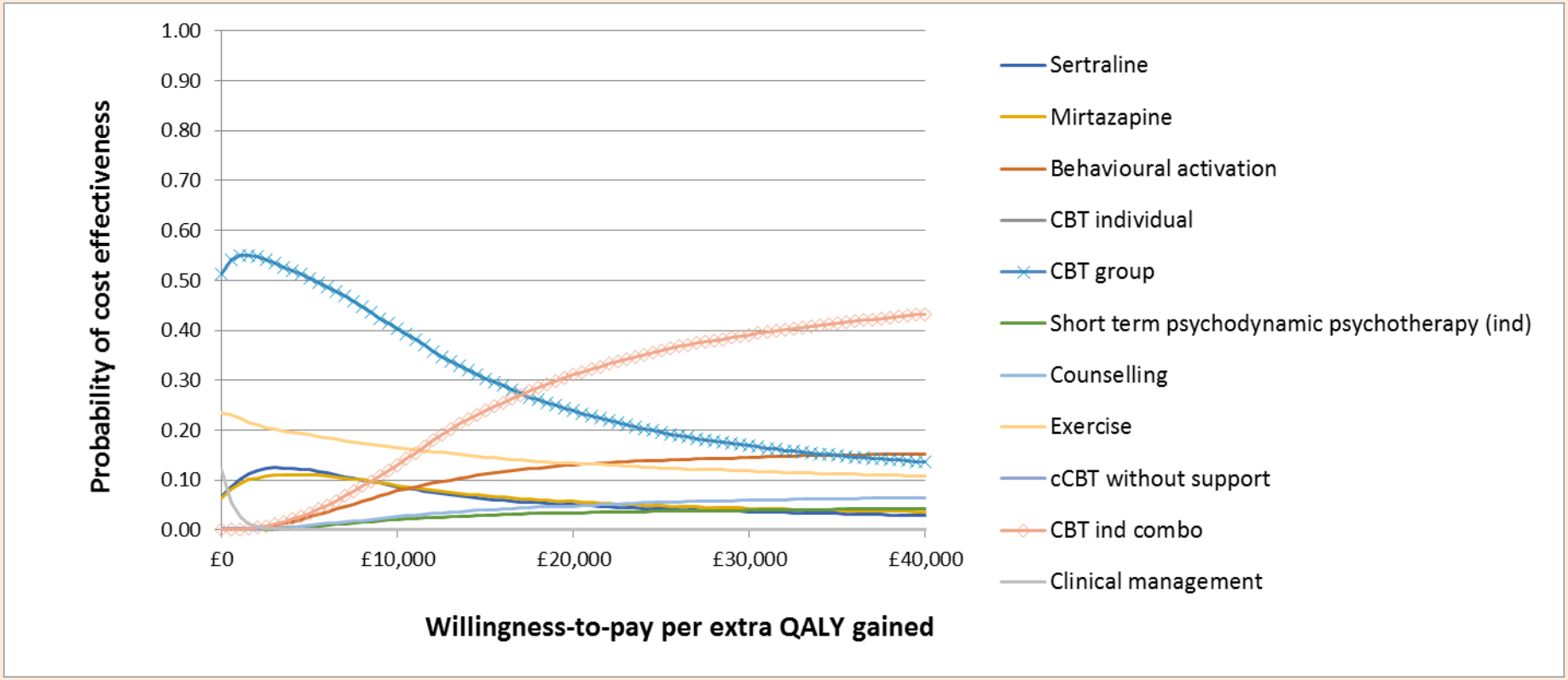
Acute treatment option	Incremental cost effectiveness (each option vs next most cost-effective option)	Probability being best[1] (step-wise approach)	Mean ranking (step-wise approach)
BA	£13,694 /QALY	0.268	3.49
Sertraline	£11,689/QALY	0.236	2.92
Physical exercise prog	Short term PDPT vs exercise £22,861/QALY	0.273	3.01
Short term PDPT	£19,595/QALY	0.241	2.70
Mirtazapine	Counselling vs mirtazapine £22,327/QALY	0.442	2.04
Counselling	£828/QALY	0.527	1.90
CBT individual	£18,557/QALY	0.423	1.92
Clinical management	Clinical management dominant	0.630	1.37
cCBT		1.000	1.00

Notes
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
BA: behavioural activation; CBT: cognitive behavioural therapy; PDPT: psychodynamic psychotherapy; prog: programme; sertr: sertraline

The CEAC and CEAF of the analysis are shown in Figure 43 and Figure 44, respectively. It can be seen that CBT group is the most cost-effective option at cost effectiveness thresholds up to £18,000/QALY, with a probability that reaches 0.55 at low cost effectiveness thresholds that are close to zero and then drops down to 0.26. For higher cost effectiveness thresholds, CBT individual combined with sertraline is the most cost-effective option for the treatment of more severe depressive episodes, with a probability of cost effectiveness that starts at 0.29 and reaches 0.43 at a cost effectiveness threshold of £40,000/QALY

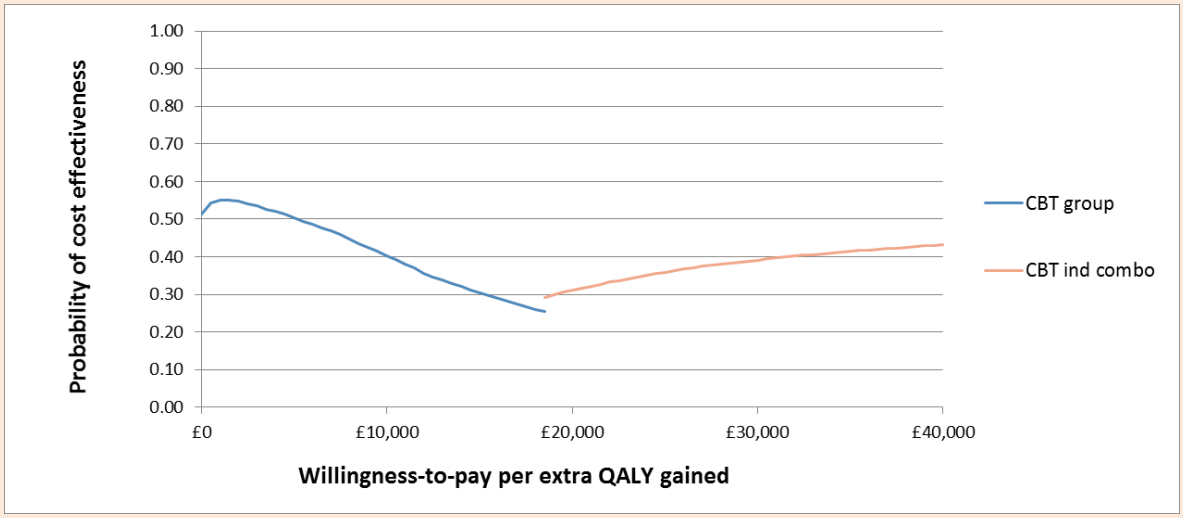
Update 2017

Figure 43. Cost-effectiveness acceptability curves of interventions for the treatment of a new episode of more severe depression in adults



Update 2017

Figure 44. Cost-effectiveness acceptability frontier of interventions for the treatment of a new episode of more severe depression in adults



Results were overall robust to alternative scenarios tested in one-way deterministic sensitivity analysis, with the following exceptions:

- When the higher utility value from Mann, Gilbody et al. (2009) was attached to more severe depression (translating into a more limited scope for HRQoL improvement following successful treatment), there were some changes in deterministic cost effectiveness ranking, which became as follows: CBT group, CBT individual combined with sertraline, BA, sertraline, mirtazapine, physical exercise programme, short term PDPT, counselling, clinical management, CBT individual, cCBT.
- When the cost of relapse was reduced by 50%, the cost effectiveness of mirtazapine improved relative to short-term PDPT and counselling; when the cost of relapse increased by 50%, the cost effectiveness of mirtazapine decreased relative to short-term PDPT and counselling.
- When all psychological interventions were assumed to be delivered by a band 5 PWP, the intervention cost of individual high-intensity psychological interventions was reduced and their relative cost effectiveness increased, resulting in changes in ranking. According to this scenario, the order of interventions from the most to the least cost-effective in deterministic analysis was as follows: CBT individual combined with sertraline, CBT group, BA, short term PDPT, counselling, sertraline, CBT individual, mirtazapine, physical exercise programme, clinical management, cCBT. Assuming that individual high-intensity psychological interventions were delivered by a band 6 therapist had a very similar impact on the results, with the only change being that CBT individual was ranked just after mirtazapine and before the physical exercise programme. It needs to be noted that combining a scenario of delivery of individual psychological interventions by a band 5 PWP with delivery of group psychological interventions by a band 7 clinical psychologist had no impact on the cost effectiveness of group CBT, which remained the second most cost-effective option.
- When counselling was assumed to be delivered in 8 sessions instead of 16, it became the 4th most cost-effective option, following BA.
- When the baseline discontinuation associated with sertraline was reduced by 20%, the cost effectiveness of mirtazapine improved relative to short-term PDPT and counselling; when the baseline discontinuation increased by 20%, the cost effectiveness of mirtazapine decreased relative to short-term PDPT and counselling.

The results of the probabilistic sensitivity analysis that utilised data on response in completers from the respective bias NMA model are shown in Table 312. Only minor

Update 2017

changes were observed in terms of effectiveness. Counselling achieved a higher ranking in terms of effectiveness and mirtazapine ranked just above the physical exercise programme, but otherwise effectiveness rankings remained the same. Regarding cost effectiveness, there were some changes in the ranking of interventions, which became as follows: CBT group (with a 0.26 probability of being cost-effective at the NICE lower cost-effectiveness threshold of £20,000/QALY), CBT individual combined with sertraline, BA, sertraline, mirtazapine, short term PDPT, counselling, physical exercise programme, clinical management, CBT individual and cCBT without or with minimal support.

Table 312: Results of economic modelling: interventions for people with a new episode of more severe depression – sensitivity analysis based on bias-adjusted NMA models (mean values from probabilistic analysis)

Acute treatment option	Mean per person			ICER (£/QALY)	NMB / person	Prob best[1]	Mean rank
	QALY	Interv cost	Total cost				
CBT individual + sertr	1.511	£1,312	£3,098	£23,955	£27,119	0.261	3.80
BA	1.472	£882	£2,743	ext domin	£26,706	0.122	5.05
Counselling	1.465	£1,195	£3,044	dominated	£26,258	0.065	6.12
Short term PDPT	1.465	£1,172	£3,023	dominated	£26,269	0.036	6.03
CBT group	1.462	£129	£1,930		£27,312	0.264	3.25
Sertraline	1.431	£100	£2,031	dominated	£26,598	0.073	4.57
CBT individual	1.425	£1,084	£3,010	dominated	£25,491	0.001	8.11
Mirtazapine	1.418	£104	£2,060	dominated	£26,298	0.057	5.67
Physical exercise prog	1.409	£94	£2,008	dominated	£26,167	0.118	6.47
Clinical management	1.380	£80	£2,008	dominated	£25,591	0.000	7.91
cCBT	1.367	£85	£2,099	dominated	£25,232	0.002	9.02

Notes:
1 at the NICE lower cost-effectiveness threshold of £20,000/QALY
BA: behavioural activation; CBT: cognitive behavioural therapy; cCBT: computerised cognitive behavioural therapy; ext domin: extendedly dominated; ICER: incremental cost effectiveness ratio; interv: intervention; NMB: net monetary benefit; PDPT: psychodynamic psychotherapy; Prob: probability; prog: programme; sertr: sertraline

Update 2017

14.4 Discussion – conclusions, strengths and limitations of economic analysis

The guideline economic analysis assessed the cost effectiveness of a range of pharmacological, psychological, physical and combined interventions for the treatment of new depressive episodes in adults with less or more severe depression treated in primary care. The interventions assessed were determined by the availability of efficacy and acceptability data obtained from the NMAs that were conducted to inform this guideline. Specific interventions were used as exemplars within each class, so that results of interventions can be extrapolated, with some caution, to other interventions of similar resource intensity within their class.

In people with less severe depression, pharmacological treatment, group psychological interventions and other low-intensity psychological and physical interventions were the most cost-effective options. These were followed by high intensity psychological interventions alone or in combination with pharmacological treatment, a number of which appeared to be less cost-effective than clinical management. The ranking of interventions, from the most to least cost-effective, was as follows: mirtazapine, CBT group, physical exercise programme, citalopram (representing SSRIs), cCBT with support (representing self-help with support),

physical exercise programme combined with sertraline, psychoeducational group programme, Coping with Depression group course (representing behavioural therapies delivered in groups), cCBT without or with minimal support (representing self-help without or with minimal support), CBT individual, BA (representing individual behavioural therapies), IPT combined with citalopram, clinical management by GPs (reflecting pill placebo trial arms), IPT, short term PDPT individual, short term PDPT individual combined with citalopram (or another antidepressant), counselling, CBT individual combined with citalopram (or another antidepressant). The probability of mirtazapine being the most cost-effective option was 0.45 at the NICE lower cost-effectiveness threshold of £20,000/QALY.

In people with more severe depression, the combination of CBT individual and sertraline (or another antidepressant) appeared to be the most cost-effective option, with a probability of 0.31 at the NICE lower cost-effectiveness threshold of £20,000/QALY. This was followed by CBT group, BA (representing individual behavioural therapies), sertraline (representing SSRIs), physical exercise programme, short term PDPT individual, mirtazapine, counselling, CBT individual, clinical management by GPs (reflecting pill placebo trial arms), and, finally, cCBT without or with minimal support (representing self-help without or with minimal support) which was the least cost-effective option in this population.

Results of the economic analysis were overall robust to different scenarios explored through sensitivity analysis. Attaching higher utility values to the states of less and more severe depression, which reduced the scope for HRQoL improvement following successful treatment, resulted in a reduction in the relative cost effectiveness of high intensity psychological interventions (i.e. BA, CBT individual, counselling, IPT, short-term PDPT) alone or in combination with drugs. Consequently, the relative cost effectiveness of pharmacological treatments, group psychological therapies and other low intensity psychological and physical interventions improved. In addition, in people with less severe depression, when the cost of relapse was assumed to be 50% lower than the base-case value, IPT combined with citalopram, BA and CBT individual became less cost-effective than clinical management. In contrast, when all psychological interventions were assumed to be delivered by a band 5 PWP, the intervention cost of individual high-intensity psychological interventions was reduced, their relative cost effectiveness increased, and their rankings improved. Nevertheless, the relative effectiveness of group CBT was not affected by this scenario, even if it was assumed to be delivered by band 7 clinical psychologists (instead of band 5 PWPs that were assumed to be delivering group CBT in base-case analysis). The cost effectiveness of counselling improves if it is delivered in 8 instead of 16 sessions. In probabilistic sensitivity analyses that utilised data on response in completers from the respective NMA models adjusted for bias relating to study sample size, only small changes in ranking were observed and the top 4 cost-effective interventions for each study population remained the same.

The analysis utilised clinical effectiveness parameters derived from NMAs of 4 different outcomes (treatment discontinuation, discontinuation due to side effects in people who discontinued treatment, response in completers and remission in completers) conducted separately for each population of interest. This methodology enabled evidence synthesis from both direct and indirect comparisons between interventions, and allowed simultaneous inference on all treatments examined in pair-wise trial comparisons while respecting randomisation (Lu and Ades 2004, Caldwell, Ades et al. 2005). The quality and limitations of RCTs considered in the NMAs have unavoidably impacted on the quality of the economic model clinical input parameters. For example, economic results may be have been affected by reporting and publication bias, although bias-adjusted models and respective sensitivity analyses tested the impact of bias relating to small study size on the results of the economic analyses.

It needs to be noted that the data that informed the NMA and the economic analyses and some of the NMA outputs are characterised by limitations:

Update 2017

A number of interventions assessed in the economic analyses were informed by limited data or borrowed efficacy data from a different (though similar) type of intervention, within the same class, if possible. In less severe depression, mirtazapine data were based on a very limited number of participants for discontinuation (N=45) and response in completers (N=27) and were absent for remission in completers; for the latter outcome, citalopram data were borrowed to inform the mirtazapine arm of the model. Therefore, results for mirtazapine in the treatment of less severe depressive episodes should be interpreted with caution. Similarly, data on IPT combined with citalopram were based on a limited number of participants in the combination of IPT plus antidepressants class (N=76 for discontinuation and N=61 for response in completers). Data on response in completers for less severe depression were limited for other classes as well (N=40 for psychoeducation, N=73 for counselling, N=40 for the combination of CBT with antidepressant, and N=62 for the combination of exercise with antidepressant or CBT). Limited data on remission in completers were available for the classes of psychoeducation (N=40), counselling (N=59), the combination of CBT with antidepressant (N=53), and the combination of exercise with antidepressant (N=88).

For more severe depression, limited data on response in completers were available for the classes of counselling (N=101), behavioural therapies (N=16) and the combination of CBT with antidepressant (N=57). Remission in completers data were very limited for mirtazapine (N=49), and the combination of CBT with antidepressant class (N=31). No remission in completers data were available for BA and CBT group interventions, which borrowed data from CBT individual, for short-term PDPT and counselling, which borrowed data from IPT (N=62), and for physical exercise programme and cCBT without or with minimal support, which borrowed data from cCBT with support (N=43).

In addition, two of the NMAs that informed the economic analysis, remission in completers in less severe depression and discontinuation in more severe depression, were characterised by inconsistency between direct and indirect evidence, and therefore their results should be interpreted with caution.

The limitations characterising the data included in the NMAs and the NMA outputs informing the economic analyses should be considered when interpreting the cost effectiveness results.

Baseline risks (discontinuation, discontinuation due to intolerable side effects, response and remission) were estimated based on a review of naturalistic studies. Available data suggested that recovery over time is characterised by a Weibull distribution, in which the events rates are proportional to a power of time. Estimation of the distribution parameters determined the probability of response and remission at 12 weeks for both less and more severe depression, based on a study that provided relevant data specific to different levels of depressive symptom severity.

The time horizon of the analysis was 12 weeks of acute treatment plus 2 years of follow up, which included maintenance treatment, as appropriate, for people who remitted following successful acute treatment. This time horizon was considered adequate to capture the full costs and effects of a course of treatment for depression (including acute and, if appropriate, maintenance treatment).

Utility data used in the economic model were derived from a systematic review of studies reporting utility data for depression-related health states that were generated using the EQ-5D and the UK population tariff, as recommended by NICE.

Intervention costs were estimated based on relevant information provided in the studies included in the NMA supplemented by GC expert opinion, in order to reflect routine NHS practice. NHS and PSS costs incurred by adults with depression following remission, treatment discontinuation, lack of adequate response or relapse were derived from a large

Update 2017

(N=88,935) naturalistic study that aimed to estimate health service use and costs associated with non-remission in people with depression using data from a large primary care UK general practice research database (Byford, Barrett et al. 2011). Resource estimates and unit costs were updated with 2016 cost data and supplemented with further evidence according to GC expert advice, where appropriate, to reflect current routine practice in the UK NHS.

The impact of intolerable side effects that led to treatment discontinuation as well as of other common side effects of pharmacological or combined treatments on HRQoL and costs associated with their management was incorporated in the economic analysis. No side effects were considered for people receiving non-pharmacological interventions; however, people receiving non-pharmacological treatments for depression are also expected to experience a range of events such as headaches, nausea or vomiting, etc. Therefore, the economic analysis may have overestimated the impact of common side effects from antidepressants relative to other treatments and thus underestimated their relative cost effectiveness. On the other hand, other less common side effects associated with treatment with antidepressants (such as upper gastrointestinal bleeds and falls) were not considered in the economic model. Such side effects result in considerable reduction in HRQoL and high costs for their management; nevertheless, they are relatively rare and therefore their omission is unlikely to have significantly impacted on the model results, although it is acknowledged as a limitation that has potentially overestimated the cost effectiveness of drugs or combined interventions with a drug component relative to other interventions.

14.5 Overall conclusions from the guideline economic analysis

In people with less severe depression, pharmacological treatment, group psychological interventions and other low-intensity psychological and physical interventions were the most cost-effective options. These were followed by high intensity psychological interventions alone or in combination with pharmacological treatment, a number of which appeared to be less cost-effective than clinical management. The ranking of interventions, from the most to least cost-effective, was as follows: mirtazapine, CBT group, physical exercise programme, citalopram (representing SSRIs), cCBT with support (representing self-help with support), physical exercise programme combined with sertraline, psychoeducational group programme, Coping with Depression group course (representing behavioural therapies delivered in groups), cCBT without or with minimal support (representing self-help without or with minimal support), CBT individual, BA (representing individual behavioural therapies), IPT combined with citalopram, clinical management by GPs (reflecting pill placebo trial arms), IPT, short term PDPT individual, short term PDPT individual combined with citalopram (or another antidepressant), counselling, CBT individual combined with citalopram (or another antidepressant). The probability of mirtazapine being the most cost-effective option was 0.45 at the NICE lower cost-effectiveness threshold of £20,000/QALY.

In people with more severe depression, the combination of CBT individual and sertraline (or another antidepressant) appeared to be the most cost-effective option, with a probability of 0.31 at the NICE lower cost-effectiveness threshold of £20,000/QALY. This was followed by CBT group, BA (representing individual behavioural therapies), sertraline (representing SSRIs), physical exercise programme, short term PDPT individual, mirtazapine, counselling, CBT individual, clinical management by GPs (reflecting pill placebo trial arms), and, finally, cCBT without or with minimal support (representing self-help without or with minimal support), which was the least cost-effective option in this population.

In both populations, the relative cost effectiveness of high intensity psychological interventions, alone or combined with antidepressants, improves when these are delivered by less specialised therapists, such as Band 5 PWPs or Band 6 therapists (instead of Band 7 clinical psychologists) and deteriorates when higher utility values are assumed at baseline, as the scope for HRQoL improvement following successful treatment is more limited. In

Update 2017

people with less severe depression the relative cost effectiveness of individual high-intensity psychological therapies is reduced when a 50% lower cost of relapse is assumed at baseline. The cost effectiveness of counselling improves if it is delivered in 8 instead of 16 sessions.

Conclusions from the guideline economic analysis refer mainly to people with depression who are treated in primary care for a new depressive episode; however, they may be relevant to people in secondary care as well, given that clinical evidence was derived from a mixture of primary and secondary care settings (however, it needs to be noted that costs utilised in the guideline economic model were mostly relevant to primary care).

Results need to be interpreted with caution due to the limited evidence base characterising some of the interventions assessed in the models and methodological limitations characterising some of the NMAs that were used to populate the economic analyses.

15 Abbreviations

3MSE	Modified Mini-Mental State Examination
5-HT	5-hydroxytryptymine
A&E	Accident and Emergency Department
ACT	acceptance and commitment therapy
AD	antidepressant
ADI	Amritsar Depression Inventory
ADM	antidepressant medication
ADQ	average daily quantities
AfC	Agenda for Change
AGREE	Appraisal of Guidelines for Research and Evaluation Instrument
AMED	Allied and Alternative Medicine Database
AMI	autobiographical memory impairment
AMS	amisulpride
AP	antipsychotic
APA	American Psychiatric Association
APNR	acute phase non-responders
ASEX	Arizona Sexual Experience scale
AUC	area under the curve
b.i.d.	twice a day
BA	behavioural activation
BABCP	British Association for Behavioural and Cognitive Psychotherapies
BAC	British Association for Counselling
BACP	British Association for Counselling and Psychotherapy
BAI	Beck Anxiety Inventory
BASDEC	Brief Assessment Schedule Depression Cards
BD	bipolar disorder
BDQ	brief disability questionnaire
BDI	Beck Depression Inventory
BDT	brief dynamic therapy
BIDS	Brief Inventory for Depressive Symptoms
BLIPS	Brief Limited Intermittent Psychotic Symptoms

Update 2017

BLRI	Barrett-Lennard Relationship Inventory
BME	black and minority ethnic
BMQ	Beliefs about Medicines Questionnaire
BMT	behavioural marital therapy
BOCF	baseline observation carried forward
BPD	borderline personality disorder
BPI	brief pain inventory
BPIT	brief psychodynamic-interpersonal therapy
Bpn	bupropion XL
BPRS	Brief Psychiatric Rating Scale
BSP/BS	brief supportive psychotherapy
BT	behaviour therapy
BtB	Beating the Blues
BZD	benzodiazepine
C	completers analysis
CADET	Collaborative Depression Trial
CAGE	a short assessment for alcohol misuse
CARE	Comprehensive Assessment and Referral Evaluation
CAT	Cliet Assessment of Treatment
CAT	cognitive analytic therapy
CAU	care as usual
CBASP	cognitive behavioural analysis system of psychotherapy
C-BDI	Chinese Beck Depression Inventory
CBT	cognitive behavioural therapy
CCBT/cCBT	computerised cognitive behavioural therapy
CCC	clinical classification categories
CCDAN	Cochrane Centre for Depression, Anxiety and Neurosis
CCG	Clinical Commissioning Group
CCSS	Caribbean Culture-Specific Screen for emotional disorders
CCT	client-centred treatment
CDRS-SR	Carroll Depression Rating Scale (Self-Report)
CDS	Chronic Disease Score

Update 2017

CEAC	cost-effectiveness acceptability curve
CEAF	cost-effectiveness acceptability frontier
CEEG	continuous electroencephalography
CES-D	Centre of Epidemiology Studies – Depression
CFB	change from baseline
CGI	Clinical Global Impressions
CI	confidence interval
CIDI (-SF)	Composite International Diagnostic Interview (-Short Form)
CIGP-CD	Cognitive-Interpersonal Group Psychotherapy for Chronic Depression
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CIS-R	Clinical Interview Schedule-Revised
Cit/cital	citalopram
clr	cluster randomised (adjusted)
CM	care management/clinical management
CMB	combined
CMBN	combined arms
CMHN	community mental health nurse
CMHT	community mental health team
CNS	central nervous system
CNSLNG	counselling
Cntl	control
CNTRL	control
COMB	combination of 12 weeks' antidepressant treatment and 16 sessions of CBT with 6 months' maintenance therapy and 6 months' follow-up (Strategy B in this guideline)
Combo	combined treatment (used in the Appendices only)
COPE	Calendar of Premenstrual Experiences
CORE	Centre for Outcomes, Research and Effectiveness
CORE (-OM)	Clinical Outcomes in Routine Evaluation (-Outcome Measure)
CPA	Care Programme Approach
CPN	community psychiatric nurse
CPRS	Comprehensive Psychopathological Rating Scale
C-R	clinician-reported

Update 2017

CRHT	crisis resolution and home treatment
CRHTT	crisis resolution and home treatment team
CrI	credible interval
CSPRS	Collaborative Study Psychotherapy Rating Scale
CSQ (-8)	Client Satisfaction Questionnaire (-8 items)
CT	cognitive therapy
Ctp	citalopram
CTS	Cognitive Therapy Scale
CWD	Coping with Depression
D	dysthymia
DA	dopamine
DAI	Drug Attitude Index
DALY	disability adjusted life years
DBM	demineralised bone matrix
DBS	deep brain stimulation
DESS	Discontinuation Emergent Signs and Symptoms
df	degrees of freedom
DIC	deviance information criterion
DIS	Diagnostic Interview Schedule
DOI	declaration of interests
DP	day patient
DPDS	depression subscale of the Short-CARE
DRP (-PC)	Depression Recurrence Prevention Program (-psychiatric consultation)
DSM (–II, –III, –IV, –TR, –R)	Diagnostic and Statistical Manual of Mental Disorders of the American Psychiatric Association (2nd edition, 3rd edition, 4th edition, Text Revision, Revision)
Dsp	desipramine
dul/dulox	duloxetine
ECG	electrocardiogram
ECT	electroconvulsive therapy
EDS	Edinburgh Depression Scale
EED	Economic Evaluation Database
EEG	electroencephalogram

Update 2017

EFT	emotion-focused therapy
EMBASE	Excerpta Medica Database
EQ-5D	European Quality of Life-5 Dimensions
ER	extended release
ERIC	Education Resources Information Center
Escit/esc	escitalopram
EuroQOL	European Quality of Life
F	female
FDA	US Food and Drug Administration
Flp	flupenthixol
FLU/fluox/flx/flu	fluoxetine
Flv/Fvx	fluvoxamine
G	group
GAD	generalised anxiety disorder
GAF	Global Assessment of Functioning
GAS	Global Assessment Scale
GBP	British pounds sterling
GC	Guideline Committee
gCBT	group cognitive behavioural therapy
GDG	Guideline Development Group
GDS	Geriatric Depression Scale
GHC	Group Health Cooperative
GHQ	General Health Questionnaire
GMS-AGECAT	Geriatric Mental State-Automated Geriatric Examination for Computer Assisted Taxonomy
GP	general practitioner
GPc	general practitioner care
GPRD	General Practice Research Database
GPT	group psychotherapy
GRADE	Grading of Recommendations Assessment, Development and Evaluation
GRP	Guideline Review Panel
GSDS	Groningen Social Disabilities Schedule

GSH	guided self-help
GSS	Global Seasonality Score
HADS	Hospital Anxiety and Depression Scale
HADS (-D)	Hospital Anxiety and Depression Scale (-Depression)
HAM-A	Hamilton Anxiety Rating Scale
HAMD/HAM-D	Hamilton Depression Rating Scale
HAP	Human Activities Profile
HAQ	health assessment questionnaire
HCl	hydrochloride
HIRU	Health Information Research Unit
HLM	hierarchical linear modelling
HMIC	Health Management Information Consortium
HMO	health maintenance organisation
HMSO	Her Majesty's Stationery Office
HMU	head-mounted unit
HRQoL	health-related quality of life
HRSD	Hamilton Rating Scale for Depression
HRT	hormone replacement therapy
HSCIC	Health and Social Care Information Centre
HSCL	Hopkins Symptom Checklist
HTA	health technology assessment
IAPT	Improving Access to Psychological Therapies
ICC	intracluster correlation coefficient
ICD (-9, -10)	International Classification of Diseases (9th revision; 10th revision)
ICD-10	ICD–10 Classification of Mental and Behavioural Disorders
ICER	incremental cost-effectiveness ratio
ICM	imipramine + clinical management
ICSD-2	International Classification of Sleep Disorders-2
ICT	integrative cognitive therapy
IDS	Inventory for Depressive Symptomatology
IHD	ischaemic heart disease
Imp	imipramine

Update 2017

IMPACT	a collaborative care for depression programme at the University of Washington
Int	intervention
Ip	interpersonal therapy for dysthymic disorder
IP	Inpatient
IPD	interpersonal difficulties
IPT	interpersonal therapy
IPT (-M, -D)	interpersonal therapy (-maintenance, -for dysthymia)
ITT	intention to treat
K	number of studies
K10	Kessler-10
KPDS	Kleinian Psychoanalytic Diagnostic Scale
LD3	low dose (three times per week)
LD5	low dose (five times per week)
LED	light-emitting diode
LGBT	lesbian, gay, bisexual and transgender
Li	lithium
LOCF	last observation carried forward
LOF	lofepramine
LOR	log-odds ratio
LR-	negative likelihood ratio
LR+	positive likelihood ratio
LSP	Life Skills Profile
LVCF	last value carried forward
M	male
MADRS	Montgomery–Åsberg Depression Rating Scale
MAJOR	major depression arm of study
MANSA	Manchester short assessment of quality of life
MAOI	monoamine-oxidase inhibitor
MBCBT	mindfulness-based CBT
MBCT	mindfulness-based cognitive therapy
MBSR	mindfulness-based stress reduction
mcl	moclobemide

MD	mean difference/major depression
MDD	major depressive disorder
MEDLINE	Medical Literature Analysis and Retrieval System Online
MHI (-5)	Mental Health Inventory (-5 items)
MHRA	Medicines and Healthcare products Regulatory Agency
MHT	mental health team
MI	myocardial infarction
MIDAS	Module for Meta-analytical Integration of Diagnostic Test Accuracy Studies
MINI	Mini International Neuropsychiatric Interview
MINOR	minor depression arm of study
MMPI	Minnesota Multiphasic Personality Inventory
MMQ	Maudsley Marital Questionnaire
MMRM	Mixed-Effect Model Repeated Measure
MMSE	Mini-Mental State Examination
Mnp	minaprine
MOS-SF-20	Medical Outcomes Study-Short Form-20 items
MPS	Maier and Philipp (core mood stability) Subscale
Mpt	maprotiline
MRC	Medical Research Council
MSE	Mental State Examination
MSQ	Mental Status Questionnaire
n	number of participants
N	total number of participants
N/A	not applicable
N/n	number of participants
N/R	not reported
NA	noradrenaline
NA	not available
NARI	noradrenaline reuptake inhibitor
NaSSA	noradrenaline and specific serotonin antidepressant
NCC	National Collaborating Centre
NCCMH	National Collaborating Centre for Mental Health

ND	non-directive
NEF	nefazodone
NEO (-FFI)	NEO Personality Inventory (-Five-Factor Inventory)
NGA	National Guideline Alliance
NHS	National Health Service
NICE	National Institute for Clinical Excellence (before April 2005) National Institute for Health and Clinical Excellence (from 1 April 2005 to 31 March 2013) National Institute for Health and Care Excellence (from 1 April 2013)
NIMH	National Institute of Mental Health
nm	nanometres
NMA	network meta-analysis
NMB	net monetary benefit
NNH	number needed to harm
NNT	number needed to treat
Nort	nortriptyline
NOS	not otherwise specified
NPV	negative predictive value
NR	not reported
NSAID	non-steroidal anti-inflammatory drug
NSF	National Service Framework
OCD	obsessive-compulsive disorder
OHE HEED	Office of Health Economics Health Economic Evaluations Database
OIS	optimal information size
Olz	olanzapine
ONS	Office for National Statistics
OpenSIGLE	system for information on Grey Literature in Europe
OR	odds ratio
OT	occupational therapy/therapist
Parox/prx/px	paroxetine
PARQ	Physical Activity Readiness Questionnaire
PASE	Physical Activity Scale for the Elderly
PC	personal computer
PCA	Prescription Cost Analysis

P-CM	placebo + clinical management
PCMHW	primary care mental health worker
PCP	primary care practitioner
PCT	Primary Care Trust
PD	personality disorder
PDPT	psychodynamic psychotherapy
PDAS	psychotic depression assessment scale
PE	process experiential treatment
PEP (+PC)	psychoeducational prevention programme (+psychiatric consultation)
PF-SOC	Problem-Focused Style of Coping scale
PGEM	pharmacist guided education and monitoring
PGI	Patient Global Impression scale
PGMS	Philadelphia Geriatric Morale Scale
PHD3	public health dose (180 minutes of moderate-intensity exercise per week, three times per week)
PHD5	public health dose (180 minutes of moderate-intensity exercise per week, five times per week)
PHQ	Patient Health Questionnaire
PHQ (-9)	Patient Health Questionnaire (-9 items)
Phz	phenelzine
PICO	population intervention comparison outcome
PLA/Plb/pbo/pb	placebo
POMS	Profile of Mood States
PP	psychodynamic psychotherapy
PR interval	the part of the electrocardiogram between the beginning of the P wave (atrial depolarisation) and the QRS complex (ventricular depolarisation)
PRIME-MD	Primary Care Evaluation of Mental Disorders
PRT	progressive resistance training
PS	problem solving
PSE	Present State Examination
PSS	personal social services
PSSRU	Personal Social Services Research Unit
PST/PS (PC)	problem-solving therapy (-primary care)

Update 2017

PsycINFO	Psychological Information Database
Pt/s	patient/s
PTSD	post-traumatic stress disorder
PWP	psychological wellbeing practitioner
QALM	quality-adjusted life month
QALY	quality-adjusted life year
QI	quality improvement
QIDS-SR	Quick Inventory of Depressive Symptomatology-Self Report
QLDS	Quality of Life Depression Scale
QoL	Quality of Life
QoLI	Quality of Life Inventory
QRS interval	period from the start of the Q wave to the end of the S wave (time for ventricular depolarisation)
QT interval	period from the start of the Q wave to the end of the T wave (duration of ventricular electrical activity)
QTc	corrected QT interval
QWB-SA	Quality of Well-Being Scale
RAND-36	A 36-item health survey by RAND
RANLab	Random Agent Networks model application
RCT	randomised controlled trial
RD	risk difference
RDC	Research Diagnostic Criteria
REBT	rational emotive behaviour therapy
RET	rational emotive therapy
RFCBT	rumination-focused CBT
RIMA	reversible inhibitors of monoamine oxidase
ROB	risk of bias
ROC	receiver operator characteristic
RQ	review question
RR	relative risk/risk ratio
RS	rating scale
RSMD	Rating Scale for Mania and Depression
rTMS	repetitive transcranial magnetic stimulation

Update 2017

Rts	ritanserin
SAD	seasonal affective disorder
SAS	Spielberger State/Trait Anxiety Scale
SAS-M	Social Adjustment Scale-modified
SAS-SR	Social Adjustment Scale-Self Report
SASS	Social Adaptation Self-evaluation Scale
SC	standard care
SCID (-IV, -PQ)	Structured Clinical Interview for DSM (-IV, -Personality Questionnaire)
SCL (-20, -90, -R)	Symptom Checklist (-20 items, -90 items, -Revised)
SD	standard deviation
SDS	Sheehan Disability Scale
SE	side effects
SE	standard error
SEM	standard error of the mean
SF-12, -36	12-/36-item short form health survey
SFS	Social Functioning Scale
SFX	significant effects
SG	standard gamble
Short-CARE	Comprehensive Assessment Referral Evaluation (short) SIGH (-SAD, -SR) Structured Interview Guide for the Hamilton Depression Rating Scale (-Seasonal Affective Disorders, -Self Rating)
SIGN	Scottish Intercollegiate Guidelines Network
SJW	St John's wort
SMD	standardised mean difference
SNRI	serotonin–noradrenaline reuptake inhibitor
SOFAS	Social and Occupational Functioning Assessment Scale
SPC	Summary of Product Characteristics
SPSP	short psychodynamic supportive psychotherapy
SQ-SS	Symptom Questionnaire-Somatic Subscale
SR	sustained release
S-R	self-reported
SRT	social rhythm therapy
Srtl/stl/st	sertraline

Update 2017

SSRI	selective serotonin reuptake inhibitor
STAI	State-Trait Anxiety Inventory
STAR*D	Sequenced Treatment Alternatives to Relieve Depression
STPT	short-term psychodynamic psychotherapy
t.i.d	three times a day
T1	end of trial
T2	6 months after end of trial
T3	triiodothyronine
TA	technology appraisal
TAU	treatment as usual
TCA	tricyclic antidepressant
TCM (-TP)	telephone care management (-telephone psychotherapy)
TDCRP	NIMH Treatment of Depression Collaborative Research Programme
tDCS	transcranial direct current stimulation
TDM	telephone disease management programme
TeCA	tetracyclic antidepressant
TMS	transcranial magnetic stimulation
TRD	treatment-resistant depression
TSU	NICE Guidelines Technical Support Unit
TTO	time trade-off
UC	usual care
UKCP	United Kingdom Council for Psychotherapy
VAMC	Veterans Affairs Medical Center
VAS	Visual Analogue Scale
VAX	virtual address eXtension
Ven/vfx	venlafaxine
VNS	vagus nerve stimulation
vrbl	verbal
WFSBP	World Federation of Societies of Biological Psychiatry
WHO	World Health Organization
WHOQOL	World Health Organization Quality of Life Assessment
WL/WLC	waitlist/waitlist control

Update 2017

WMD	weighted mean differences
WSAS	Work and Social Adjustment Scale
WSDS	Work and Social Disability Scale
XL/XR	extended release
YLD	years lived with disability

Update 2017

16 References

This chapter forms a separate document.

Update 2017

17 Network meta-analysis — detailed methods and results

This chapter forms a separate document.

Appendices

The following appendices are provided as separate documents:

Content	Appendix
Scope for the development of the clinical guideline	Appendix A
Declarations of interests by Guideline Committee members	Appendix B
Special advisers to the Guideline Committee	Appendix C
Stakeholders	Appendix D
Researchers contacted to request information about unpublished or soon-to-be published studies	Appendix E
Review questions and review protocols	Appendix F
Research recommendations	Appendix G
Search strategies – clinical evidence	Appendix H
Search strategies – economic evidence	Appendix I
Study characteristics	Appendix J
Clinical evidence – flow charts	Appendix K
Clinical evidence – GRADE evidence profiles	Appendix L
Clinical evidence – forest plots	Appendix M
Clinical evidence – network meta-analysis: bias adjustment methods and results	Appendix N
Economic evidence – flow chart	Appendix O
Economic evidence – health economic checklists	Appendix P
Economic evidence – evidence tables	Appendix Q
Economic evidence – economic profiles	Appendix R
Economic evidence – list of excluded studies	Appendix S
Studies and data included in the network meta-analyses	Appendix T
Study references from 2004 and 2009 guidelines	Appendix U
Superseded text from CG90 guideline	Appendix V
Network meta-analysis inconsistency checks - direct vs NMA results	Appendix W

Update 2017