Contents

Review Protocols

Service Delivery: RQ 1.1 (service delivery models)

Topic	Organisation and delivery of services
Review question	RQ.1.1 For adults with depression, what are the relative benefits and harms associated with different models for the coordination and delivery of services?
Objectives	To identify the optimal model of delivery of services for adults with an acute episode of depression, or adults whose depression has responded fully or partially to treatment.
Population	• Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on validated scales (and including those with subthreshold [just below threshold] depressive symptoms) For studies on relapse prevention: • Adults whose depression has responded to treatment (in full or partial remission) according to DSM, ICD or similar criteria, or indicated by below clinical threshold depression symptom scores on validated scales If some, but not all, of a study's participants are eligible for the review, for instance, mixed anxiety and depression diagnoses, then we will include a study if at least 80% of its participants are eligible for this review
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Intervention	Models for the coordination and delivery of services: • Collaborative care (simple and complex) • Stepped care • Medication management • Attached professional model • Care coordination • Integrated care pathways (including primary care liaison or shared care) • Measurement-based care
Comparison	• Treatment as usual • Waitlist • Any other service delivery model

Outcomes	**Critical outcomes**: • Depression symptomology (mean endpoint score or change in depression score from baseline) • Response (usually defined as at least 50% improvement from the baseline score on a depression scale) • Remission (usually defined as a score below clinical threshold on a depression scale) • Relapse (number of people who returned to a depressive episode whilst in remission) The following depression scales will be included in the following hierarchy: • MADRS • HAMD • QIDS • PHQ • CGI (for dichotomous outcomes only) • CES-D • BDI • HADS-D (depression subscale) **Important but not critical outcomes:** • Antidepressant use • Discontinuation due to any reason Outcomes will be assessed at 6 months and 12 months.
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried forward. No restriction on date for the updated search, studies published between database inception and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm Studies with <50% completion data (drop out of >50%) will be excluded.
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Coding Strategy** For this review, a coding system for classifying the complexity and type of service delivery model has been developed specifically for the purpose of this guideline. The service delivery model described in each study will be rated on this 17-

	item coding system which will generate an overall rating between 0-20 (see Table 1). Service delivery models which score above 6 will be considered a collaborative care intervention; those scoring 13+ will be coded as complex collaborative care and those scoring 6-12 will be coded as simple collaborative care. Service delivery models that score below 6 will be classified as an alternative service delivery model (e.g. care coordination) or a stand-alone psychological intervention (e.g. self-help with support). **Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** A meta-analysis using a random-effects model will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is

imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.

Table 1. Coding system for service delivery models

Collaborative Care Component Score Method

Item	Score
1. Active and integrated case recognition/identification* (Systematic identification- from a clinical database or screened positive for depression)	0 1
2. Collaborative assessment and plan included (Collaborative assessment with the patient)	0 1
3. Case Management (Case manager present- can include pharmacist for medication management)	0 1
4. Active liaison with primary care and other services (System set up for structured liaison/ regular meetings)	0 1
5. Case Manager has MH background (A prior mental health background, not just training in mental health)	0 1
6. Supervision provided for case manager	0 1
7. Senior MH professional consultation/involvement (Broad definition- just need to be available)	0 1
8. Psychoeducation delivered	0 1
9. Algorithm(s) used to determine care*	0 1
10. Integration with physical health care where necessary	0 1
11. Social/psychosocial interventions provided	0 1
12. Case manager delivers intervention	0 1
13. Medication management provided	0 1
14. Routine outcome monitoring (Scheduled, using a tool)	0 1
15. Psychological interventions provided None Low intensity High intensity	 0 1 2
16. Duration of programme contact ≤6 mths 7-12mths 1year plus	 0 1 2
17. Number of sessions (F-t-F and Telephone) ≤6 sessions 6 – 12 sessions 13 + sessions	 0 1 2
Total (maximum 20)	
***Including stepped care** **Rating** <5 – not collaborative care	

	6-12 – simple collaborative care 13+ – complex collaborative care
Heterogeneity (sensitivity analysis and subgroups)	Where possible, the influence of the following subgroups will be considered: For the review of collaborative care only: • Type of collaborative care (simple vs complex) • Stepped care component included in collaborative care intervention • Case manager background • Psychological interventions delivered as part of the model of care • Number of contacts/sessions/follow-up visits provided as part of intervention (less than 13 sessions, 13+ sessions) For all reviews: • Chronic depression • Depression with coexisting personality disorder • Psychotic depression • Older adults • BME populations • Men
Notes	The GC identified one good quality systematic review of RCTs (Coventry et al., 2014) which reviewed collaborative care interventions. The review was used as a source to identify any additional eligible studies *Coventry PA, Hudson JL, Kontopantelis E, Archer J, Richards DA, et al. (2014) Characteristics of Effective Collaborative Care for Treatment of Depression: A Systematic Review and Meta-Regression of 74 Randomised Controlled Trials. PLoS ONE 9(9): e108114.* Separate reviews (if applicable) will be conducted for service delivery models which were aimed at: 1. Treating an episode of depression 2. Preventing relapse of a future episode of depression

Service Delivery: RQ 1.2 (settings for care)

Topic	**Organisation and delivery of services**
Review question	RQ.1.2 For adults with depression, what are the relative benefits and harms associated with different settings for the delivery of care?
Objectives	To identify the optimal settings for the delivery of care for adults with depression
Population	• Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on validated scales (and including those with subthreshold [just below threshold] depressive symptoms) If the evidence specific to depression is limited then the inclusion criteria may be expanded to include those with non-psychotic severe mental illness. If some, but not all, of a study's participants are eligible for the review, then we will include a study if the majority (at least 51%) of its participants are eligible for this review.
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Intervention	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	**Critical outcomes**: • Depression symptomology (mean endpoint score or change in depression score from baseline) • Response (usually defined as at least 50% improvement from the baseline score on a depression scale)

	• Remission (usually defined as a score below clinical threshold on a depression scale) • Relapse (number of people who returned to a depressive episode whilst in remission) **Important but not critical outcomes:** • Service utilisation/resource use (e.g. antidepressant use) • Psychological functioning • Social functioning • Satisfaction • Carer distress Outcomes will be assessed at endpoint and follow-up.
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried forward. No restriction on date for the updated search, studies published between database inception and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm Studies with <50% completion data (drop out of >50%) will be excluded.
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis**

	A meta-analysis using a random-effects model will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.
Heterogeneity (sensitivity analysis and subgroups)	Where possible, the influence of the following subgroups will be considered: • Chronic depression • Depression with coexisting personality disorder • Psychotic depression • Older adults
Notes	If no RCT evidence is identified that specifically addresses the following settings: primary care, and inpatient care, then indirect evidence will be considered in the form of sub-analyses of the NMA dataset (first-line treatment of depressive episodes)

Treatment of depression: RQ 2.1-2.2 (first-line treatment)

<table>
<tr><th>Topic</th><th>First-line treatment of depression</th></tr>
<tr><td>Review question</td><td>RQ. 2.1 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?

RQ. 2.2. 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?</td></tr>
<tr><td>Objectives</td><td>To identify the most effective first-line interventions for the treatment of a new episode of depression</td></tr>
<tr><td>Population</td><td>• 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 validated scales (and including those with subthreshold [just below threshold] depressive symptoms)

If some, but not all, of a study’s participants are eligible for the review, for instance, mixed anxiety and depression diagnoses, 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 2. These thresholds are derived using standardization of depression measurement crosswalk tables (Wahl et al. 2014; Rush et al. 2003; Carmody et al. 2006; Uher et al. 2008). An anchor point of 16 on the PHQ-9 was selected on the basis of alignment with the clinical judgement of the committee and eligibility criteria in published studies. 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’).

Table 2. Severity thresholds
<table>
<tr><th>Scale</th><th>Threshold</th></tr>
<tr><td>HAMD (17-item, 21-item and 24-item)</td><td>16</td></tr>
<tr><td>MADRS (10-item)</td><td>22</td></tr>
<tr><td>PHQ-9</td><td>16</td></tr>
<tr><td>BDI-I (21-item)</td><td>22</td></tr>
<tr><td>BDI-II (21-item)</td><td>30</td></tr>
<tr><td>CES-D (20-item)</td><td>36</td></tr>
<tr><td>QIDS (16-item)</td><td>12</td></tr>
<tr><td>HADS-D (7-item)</td><td>12</td></tr>
</table>
</td></tr>
</table>

Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials where more than 20% of the population have psychotic symptoms • Trials where more than 20% of the population have a coexisting personality disorder • Trials where more than 20% of the population have chronic depression (chronic depression defined as depression for at least 2 years, or persistent subthreshold symptoms [dysthymia], or double depression [an acute episode of major depressive disorder superimposed on dysthymia]) • Trials of further-line treatment • Trials of people with Seasonal Affective Disorder (SAD) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Intervention	The following interventions will be included: **Psychological interventions**: • Behavioural therapies (including behavioural activation, behavioural therapy [Lewinsohn 1976], coping with depression group) • 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 emotion-focused therapy [EFT], non-directive/supportive/ person-centred 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) • Self-help with or without support (including cognitive bibliotherapy with or without support, computerised CBT [CCBT] with or without support, computerised psychodynamic therapy with or without support) • Art therapy

	• Music therapy • Eye movement desensitization and reprocessing (EMDR) (for depression, not PTSD) **Pharmacological interventions**: To be included, pharmacological interventions needed to be licensed in the UK and in routine clinical use for the first-line treatment of depression. SSRIs • Citalopram • Escitalopram • Paroxetine • Sertraline • Fluoxetine TCAs • Amitriptyline • Clomipramine • Lofepramine • Nortriptyline • Note: To improve connectivity, imipramine will be included in the network (because it has been used as a control in many trials) however it will not be considered as part of the decision problem SNRIs • Venlafaxine • Duloxetine Other antidepressant drugs: • Mirtazapine • Trazodone Note that if necessary for 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 or physical intervention and/or combined with a psychological or physical intervention but they will not be considered as part of the decision problem. **Physical interventions**: • Acupuncture • Exercise (including yoga) • Light therapy (for depression, not SAD)

	Psychosocial interventions: • Peer support (including befriending, mentoring, and community navigators) • Mindfulness, meditation or relaxation (including mindfulness-based stress reduction [MBSR]) The following interventions are more appropriate for subgroups of adults with depression and as such will be considered only in pairwise comparisons (and not included in the NMA): • Couple interventions, including behavioural couples therapy (for people with problems in the relationship with their partner)
Comparison	• Other active intervention (must also meet inclusion criteria above) • Treatment as usual • Waitlist • No treatment • Placebo If a study compares 'intervention + TAU vs TAU alone' it will be recoded as 'intervention vs no treatment'
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), this will be analysed for those randomised and for completers • Response (usually defined as at least 50% improvement from the baseline score on a depression scale), this will be analysed for those randomised and for completers The following depression scales will be included in the following hierarchy: • MADRS • HAMD • QIDS • PHQ • CGI (for dichotomous outcomes only) • CES-D • BDI • HADS-D (depression subscale) • 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 (see above) will be adopted for continuous outcomes **Acceptability/tolerability** • Discontinuation due to side effects (for pharmacological trials) • Discontinuation due to any reason (including side effects) **Important, but not critical, outcomes**: **Quality of life** • Quality of life (as assessed with a validated scale, including the 12-item/36-item Short-Form Survey [SF-12/SF-36], 26-item short version of the World Health Organization Quality of Life assessment [WHOQOL-BREF], EuroQoL [EQ5D], Quality of Life Depression Scale [QLDS], Quality of Life Enjoyment and Satisfaction Questionnaire [Q-LES-Q], Quality of Life Inventory [QoLI], and World Health Organization 5-item Well-Being Index [WHO-5]) **Personal, social, and occupational functioning** • Global functioning (as assessed with a validated scale, including Global Assessment of Functioning [GAF], Global Assessment Scale [GAS], and Social and Occupational Functioning Assessment Scale [SOFAS]) • Functional impairment (as assessed with a validated scale, including Sheehan Disability Scale [SDS], Social Adjustment Scale [SAS], and Work and Social Adjustment Scale [WSAS]) • Sleeping difficulties (as assessed with a validated scale, including Insomnia Severity Index [ISI] and Pittsburgh Sleep Quality Index [PSQI]) • Employment (for instance, % unemployed) • Interpersonal problems (as assessed with a validated scale, including Inventory of Interpersonal Problems [IIP]) Outcomes will be assessed at endpoint and follow-up (data for all available follow-up periods of at least 1-month post-intervention will be extracted and will be grouped into

	categories for analysis, for instance, 1-3 months, 4-6 months, 7-9 months, 10-12 months, 13-18 months, 19-24 months, and >2 years).
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried forward. Studies published between 2016 and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm Studies with <50% completion data (drop out of >50%) will be excluded.
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** Pairwise comparisons (meta-analyses using random-effects models) will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Network meta-analysis (NMA) in a Bayesian framework will also be used to synthesise the data for all eligible interventions which are connected in a network of RCT comparisons. Interventions with similar effects (as determined by the committee) will be grouped into classes and class effects models will be fitted [Dias 2018]. The relative effects of the interventions within each class will be assumed to be

	distributed around a common class mean with a within-class variance, permitting the borrowing of strength across interventions within each class. Classes which do not have enough evidence to estimate within-class variability of effects (i.e., a class with just 1 or 2 interventions) will share within-class variability with similar classes (as determined by the committee) where the variance can be estimated. For example, the individual cognitive and CBT class may borrow the within-class variance from the individual behavioural therapies class. If no such similar class is identified, we will assume zero variance in classes with only 1 or 2 interventions. In addition, the attention placebo, no treatment and TAU classes will share a within-class variance. If an 'any antidepressant' class is required to connect otherwise disconnected/excluded drugs to the network (as described under Intervention topic), its within-class variance will be equal to the maximum of the SSRI and TCA within-class variances. The random class effects assumption will be assessed by comparing the fit of fixed and random class effects models, where the former assumes the intervention effects within each class are the same (i.e., no within-class variability of effects). Continuous outcomes (SMDs) will be combined with dichotomous data to estimate intervention effects, using the methods described in the Appendix. The NMA will probably be restricted to critical outcomes at endpoint due to the likelihood of a lack of connectivity in a follow-up data network or in a network for important (but not critical) outcomes. The consistency of direct and indirect evidence will be assessed by fitting and comparing the fit of the NMA and unrelated mean effects (UME) models, the latter of which is equivalent to having separate, unrelated, meta-analyses for every pairwise contrast [Dias 2011]. Each data point's contribution to the posterior mean residual deviance for the NMA model will be plotted against that for the UME model, to visually assess if specific data points are contributing to inconsistency. If the UME suggests there is evidence of inconsistency, node-split models will be fitted to assist in identifying loops of evidence with inconsistency [Dias 2010]. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20%

	and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.
Heterogeneity (sensitivity analysis and subgroups)	Where possible, the influence of the following subgroups will be considered: • Primary care compared to secondary care • Inpatient compared to outpatient settings • Older adults (60 years and older) compared to younger adults (younger than 60 years) • BME populations • Men If the network structure allows, sensitivity analyses will be considered for depression symptoms (SMD, the primary outcome for the clinical analysis) and discontinuation for any reason and response in completers (the main outcomes for economic analysis), as follows: • Risk of bias as reflected by publication bias and study size using methods described in [Dias 2010]. We will assume possible bias in comparisons of active interventions vs inactive control and no bias between inactive control comparisons, as well as active intervention comparisons, except in comparisons where counselling is the control intervention (in which case bias against counselling will be assumed) • Validity of transitivity assumption will be explored by sensitivity analysis on SMD outcome that includes non-pharmacological trials only and examines any differences in magnitude of effects and ranking of non-pharmacological interventions compared to results from the mixed psychological, psychosocial, pharmacological and physical model

	Threshold analysis will be performed to assess the robustness of intervention recommendations due to bias [Phillippo 2018].
Notes	For interventions in the NMA it is assumed that any patient that meets all inclusion criteria is, in principle, equally likely to be randomised to any of the interventions in the synthesis comparator set. For defining routine usage of drugs, the national prescription cost data for England in 2017 - the most recent year for which relevant data existed - (Prescribing & Medicines Team, Health and Social Care Information Centre, 2017) was used. If a drug appeared in the top 15 it was included, with the exception of dosulepin which the BNF indicates should be initiated by a specialist. Cipriani et al. (2018) network meta-analysis will be used as a source for studies and data. References for crosswalk tables: Carmody, T. J., Rush, A. J., Bernstein, I., Warden, D., Brannan, S., Burnham, D., ... & Trivedi, M. H. (2006). The Montgomery Äsberg and the Hamilton ratings of depression: a comparison of measures. European Neuropsychopharmacology, 16(8), 601-611. Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., ... & Thase, M. E. (2003). The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biological psychiatry, 54(5), 573-583. Uher, R., Farmer, A., Maier, W., Rietschel, M., Hauser, J., Marusic, A., ... & Henigsberg, N. (2008). Measuring depression: comparison and integration of three scales in the GENDEP study. Psychological medicine, 38(2), 289-300. Wahl, I., Löwe, B., Bjorner, J. B., Fischer, F., Langs, G., Voderholzer, U., ... & Rose, M. (2014). Standardization of depression measurement: a common metric was developed for 11 self-report depression measures. Journal of clinical epidemiology, 67(1), 73-86. Assuming a normal distribution and using baseline mean and standard deviation data, we will explore the categorisation of less and more severe, including the percentage of studies 'definitely' within the correct category (≥70% of the study sample above cut-off) in order to aid the committee in interpreting the results.

	References for data analysis: Dias, S., Ades, A.E., Welton, N.J., Jansen, J.P., Sutton, A.J. (2018). Network meta-analysis for decision making. Hoboken, NJ: Wiley. Dias, S., Welton, N.J., Sutton, A.J., Caldwell, D.M., … & Ades, A.E. (2011). NICE DSU Technical Support Document 4: Inconsistency in networks of evidence based on randomised controlled trials. Dias, S., Welton, N.J., Caldwell, D.M., Ades A.E. (2010a). Checking consistency in mixed treatment comparison meta-analysis. Statistics in Medicine, 29(7-8), 932-44. References for heterogeneity: Dias, S., Welton, N.J., Marinho, V.C.C., Salanti, G., … & Ades A.E. (2010b). Estimation and adjustment of bias in randomised evidence by using mixed treatment comparison meta-analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society), 173(3), 613-29. Phillippo, D.M., Welton, N.J., Dias, S., Didelez, V., Ades A.E. (2018). Sensitivity of treatment recommendations to bias in network meta-analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society), 181(3), 843-67.

Treatment of depression: RQ 2.3 (relapse prevention)

Topic	Relapse prevention
Review question	RQ. 2.3 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)?
Objectives	To identify the most effective interventions for preventing relapse of depression in adults who have responded fully or partially to treatment
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 full or partial remission. If some, but not all, of a study's participants are eligible for the review, for instance, mixed anxiety and depression diagnoses, then we will include a study if at least 80% of its participants are eligible for this review.
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials where more than 20% of the population have psychotic symptoms • Trials where more than 20% of the population have a coexisting personality disorder • Trials where more than 20% of the population have chronic depression • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes) • Trials where participants are not randomised to a relapse prevention intervention following response to initial treatment e.g. continuation trials
Intervention	Interventions will be included either alone or in combination. **Psychological interventions** • Behavioural therapies (including behavioural activation, behavioural therapy [Lewinsohn 1976], coping with depression group) • Cognitive and cognitive behavioural therapies (including CBT individual or group, problem solving, rational emotive behaviour therapy [REBT], third-wave cognitive

	therapies, and mindfulness-based cognitive therapy [MBCT]) • Counselling (including emotion-focused therapy [EFT], non-directive/supportive/ person-centred counselling and relational client-centred therapy) • Interpersonal psychotherapy (IPT) • Psychodynamic psychotherapies (including short-term psychodynamic psychotherapy, long-term psychodynamic psychotherapy and psychodynamic counselling) • Psychoeducational interventions (including psychoeducational group programmes) • Self-help with or without support (including cognitive bibliotherapy with or without support, computerised CBT [CCBT] with or without support, computerised psychodynamic therapy with or without support) • Art therapy • Music therapy • Eye movement desensitization and reprocessing (EMDR) (for depression, not PTSD) **Pharmacological interventions** • SSRIs (including paroxetine, sertraline, fluoxetine, escitalopram, citalopram, fluvoxamine) • TCAs (including amitriptyline, dothiepin, imipramine, nortriptyline) • SNRIs (including duloxetine, venlafaxine, desvenlafaxine) • Mirtazapine • Antipsychotics (including olanzapine, risperidone, quetiapine)[1] • Lithium **Physical interventions** • Acupuncture • Exercise • Yoga • ECT • Light therapy (for depression, not SAD) **Psychosocial interventions**: • Peer support (including befriending, mentoring, and community navigators) • Mindfulness, meditation or relaxation (including mindfulness-based stress reduction [MBSR])

Comparison	• Other active intervention (must also meet inclusion criteria above) • Treatment as usual • Waitlist • No treatment • Placebo
Outcomes	**Critical outcomes** • Relapse (the number of participants who relapsed) **Important, but not critical, outcomes**: **Quality of life** • Quality of life (as assessed with a validated scale, including the 12-item/36-item Short-Form Survey [SF-12/SF-36], 26-item short version of the World Health Organization Quality of Life assessment [WHOQOL-BREF], EuroQoL [EQ5D], Quality of Life Depression Scale [QLDS], Quality of Life Enjoyment and Satisfaction Questionnaire [Q-LES-Q], Quality of Life Inventory [QoLI], and World Health Organization 5-item Well-Being Index [WHO-5]) **Personal, social, and occupational functioning** • Global functioning (as assessed with a validated scale, including Global Assessment of Functioning [GAF], Global Assessment Scale [GAS], and Social and Occupational Functioning Assessment Scale [SOFAS]) • Functional impairment (as assessed with a validated scale, including Sheehan Disability Scale [SDS], Social Adjustment Scale [SAS], and Work and Social Adjustment Scale [WSAS]) • Sleeping difficulties (as assessed with a validated scale, including Insomnia Severity Index [ISI] and Pittsburgh Sleep Quality Index [PSQI]) • Employment (for instance, % unemployed) • Interpersonal problems (as assessed with a validated scale, including Inventory of Interpersonal Problems [IIP]) Outcomes will be assessed at endpoint and follow-up (data for all available follow-up periods of at least 1-month post-intervention will be extracted and will be grouped into categories for analysis, for instance, 1-3 months, 4-6 months, 7-9 months, 10-12 months, 13-18 months, 19-24 months, and >2 years).
Study design	• RCTs • Systematic reviews of RCTs

Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried forward. No restriction on date for the updated search, studies published between database inception and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm Studies with <50% completion data (drop out of >50%) will be excluded.
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** Pairwise comparisons (meta-analyses using random-effects models) will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Network meta-analysis (NMA) in a Bayesian framework will also be used to synthesise the data for all eligible interventions (which are connected to the network). The NMA will be restricted to the critical outcome of relapse. A binomial likelihood and cloglog link linear model will be used (Dias et al., 2011) to allow estimation of hazard ratios between all pairs of interventions. Where possible, different NMAs will be considered for different populations according to their risk of relapse (medium or high, defined according to the number of previous episodes) and the type of previous acute treatment they received (pharmacological, psychological or combined).

	Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is <u>not</u> imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.
Heterogeneity (sensitivity analysis and subgroups)	Where possible, the following subgroup analyses will be considered: • Type of previous acute treatment received • Risk of relapse (number of previous episodes) • Remission status (participants in partial or full remission vs full remission only) • Abrupt vs slow switch to placebo
Notes	One good quality systematic review for non-pharmacological interventions for relapse prevention was identified (Clarke et al., 2015) which was used a source of studies for the review of psychological interventions. *[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)* Dias, S., Welton, N.J., Sutton, A.J., & Ades, A.E. (2011, last updated September 2016). NICE DSU Technical Support Document 2: A Generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials.

Treatment of depression: RQ 2.4-2.5 (further-line treatment)

Topic	Further-line treatment of depression
Review question	RQ. 2.4-2.5 What are the relative benefits and harms of further-line psychological, psychosocial, pharmacological and physical interventions (alone or in combination), for adults with depression showing an inadequate response to at least one previous intervention for the current episode?
Objectives	To identify the most effective interventions for people who have had no or limited response to previous treatment(s) (for the current episode), have not tolerated previous treatment(s) (for the current episode), or have treatment-resistant depression
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), or who are defined as meeting criteria for treatment-resistant depression, and who have been randomised to the further-line interventions at the point at which they had no/inadequate/limited response If some, but not all, of a study's participants are eligible for the review, then we will include a study if at least 80% of its participants are eligible for this review
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Intervention	Interventions listed below are examples of interventions which may be included either alone or in combination: <u>Psychological interventions</u>: • Behavioural therapies (including behavioural activation, behavioural therapy [Lewinsohn 1976], coping with depression group) • Cognitive and cognitive behavioural therapies (including CBT individual or group, problem solving, rational emotive behaviour therapy [REBT], third-wave cognitive therapies, Mindfulness-based Cognitive Therapy [MBCT]

	and Cognitive Behavioural Analysis System of Psychotherapy [CBASP]) • Counselling (including emotion-focused therapy [EFT], non-directive/supportive/ person-centred counselling and relational client-centred therapy) • Interpersonal psychotherapy (IPT) • Psychodynamic psychotherapies (including short-term psychodynamic psychotherapy, long-term psychodynamic psychotherapy and psychodynamic counselling) • Psychoeducational interventions (including psychoeducational group programmes) • • Self-help with or without support (including cognitive bibliotherapy with or without support, computerised CBT [CCBT] with or without support, computerised psychodynamic therapy with or without support) • Art therapy • Music therapy • Eye movement desensitization and reprocessing (EMDR) (for depression, not PTSD) <u>Psychosocial interventions</u>: • Peer support (including befriending, mentoring, and community navigators) • Mindfulness, meditation or relaxation (including mindfulness-based stress reduction [MBSR]) <u>Pharmacological interventions</u> 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] • Mirtazepine Anticonvulsants • Lamotrigine[3] Antipsychotics • Amisulpride[3] • Aripiprazole[3] • Olanzapine[3] • Quetiapine • Risperidone[3] • Ziprasidone[2] Anxiolytics • Buspirone Stimulants • Methylphenidate[3] Other agents • Lithium • Omega-3 fatty acids • Thyroid hormone[3] <u>Physical interventions</u> • Acupuncture • ECT • Exercise • Yoga • Light therapy (for depression, not SAD) 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/psychosocial/physical intervention)
Comparison	• Other active intervention (must also meet inclusion criteria above) • Treatment as usual • Waitlist • No treatment • Placebo 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
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 (usually defined as at least 50% improvement from the baseline score on a depression scale) The following depression scales will be included in the following hierarchy: • MADRS • HAMD • QIDS • PHQ • CGI (for dichotomous outcomes only) • CES-D • BDI • HADS-D (depression subscale) • HADS (full scale) **Acceptability/tolerability** • Discontinuation due to side effects (for pharmacological trials) • Discontinuation due to any reason (including side effects) **Important, but not critical, outcomes**: **Quality of life**

	• Quality of life (as assessed with a validated scale, including the 12-item/36-item Short-Form Survey [SF-12/SF-36], 26-item short version of the World Health Organization Quality of Life assessment [WHOQOL-BREF], EuroQoL [EQ5D], Quality of Life Depression Scale [QLDS], Quality of Life Enjoyment and Satisfaction Questionnaire [Q-LES-Q], Quality of Life Inventory [QoLI], and World Health Organization 5-item Well-Being Index [WHO-5]) **Personal, social, and occupational functioning** • Global functioning (as assessed with a validated scale, including Global Assessment of Functioning [GAF], Global Assessment Scale [GAS], and Social and Occupational Functioning Assessment Scale [SOFAS]) • Functional impairment (as assessed with a validated scale, including Sheehan Disability Scale [SDS], Social Adjustment Scale [SAS], and Work and Social Adjustment Scale [WSAS]) • Sleeping difficulties (as assessed with a validated scale, including Insomnia Severity Index [ISI] and Pittsburgh Sleep Quality Index [PSQI]) • Employment (for instance, % unemployed) • Interpersonal problems (as assessed with a validated scale, including Inventory of Interpersonal Problems [IIP]) Outcomes will be assessed at endpoint and follow-up (data for all available follow-up periods of at least 1-month post-intervention will be extracted and will be grouped into categories for analysis, for instance, 1-3 months, 4-6 months, 7-9 months, 10-12 months, 13-18 months, 19-24 months, and >2 years).
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried forward. Studies published between 2016 and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm Studies with <50% completion data (drop out of >50%) will be excluded.
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Data Extraction (selection and coding)**

	Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** A meta-analysis using a random-effects model will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is <u>not</u> imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.
Heterogeneity (sensitivity analysis and subgroups)	Where possible, the influence of the following subgroups will be considered: • Psychotic depression • Depression with coexisting personality disorder

	• Chronic depression
Notes	If trials specifically recruited populations with chronic depressive symptoms they would be included in this review (as opposed to RQ 2.6) if the treatment was further-line and if they reported a critical outcome. A Cochrane review of psychological therapies for treatment-resistant depression in adults was identified (Ijaz et al., 2018) which was used a source of studies for the review of psychological interventions. *[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*

Treatment of depression: RQ 2.6 (first-line treatment or relapse prevention of chronic depression)

Topic	First-line treatment or relapse prevention of chronic depression
Review question	RQ. 2.6 For adults with chronic depression or persistent subthreshold depression symptoms what are the relative benefits and harms of first-line treatment or relapse prevention with psychological, psychosocial, pharmacological and physical interventions (alone or in combination)?
Objectives	To identify the most effective strategy for the first-line treatment or relapse prevention of chronic depression or persistent subthreshold depression symptoms
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 validated scales, for at least 2 years; persistent subthreshold symptoms (dysthymia); double depression (an acute episode of MDD superimposed on dysthymia). If some, but not all, of a study's participants are eligible for the review, then we will include a study if at least 80% of its participants are eligible for this review
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials where more than 20% of the population have psychotic symptoms • Trials where more than 20% of the population have a coexisting personality disorder • Trials of further-line treatment following no/inadequate/limited response • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Intervention	Interventions listed below are examples of interventions which may be included either alone or in combination. **Psychological interventions**: • Behavioural therapies (including behavioural activation, behavioural therapy [Lewinsohn 1976], coping with depression group) • Cognitive and cognitive behavioural therapies (including CBT individual or group, problem solving, rational

	emotive behaviour therapy [REBT], third-wave cognitive therapies, Cognitive behavioral analysis system of psychotherapy [CBASP], and Mindfulness-based Cognitive Therapy [MBCT]) • Counselling (including emotion-focused therapy [EFT], non-directive/supportive/ person-centred counselling and relational client-centred therapy) • Interpersonal psychotherapy (IPT) • Psychodynamic psychotherapies (including short-term psychodynamic psychotherapy, long-term psychodynamic psychotherapy and psychodynamic counselling) • Art therapy • Music therapy • Eye movement desensitization and reprocessing (EMDR) (for depression, not PTSD) **Psychosocial interventions:** • Peer support (including befriending, mentoring, and community navigators) • Mindfulness, meditation or relaxation (including mindfulness-based stress reduction [MBSR]) **Pharmacological interventions:** Antidepressants SSRIs • Citalopram • Escitalopram • Fluvoxamine • Fluoxetine • Paroxetine • Sertraline TCAs • Amineptine[1] • Amitriptyline • Clomipramine • Desipramine[2] • Imipramine • Lofepramine • Nortriptyline MAOIs • Phenelzine TeCAs • Mianserin SNRIs • Duloxetine • Venlafaxine

	Other antidepressant drugs • Bupropion[3] • Mirtazepine • Moclobemide • Nefazadone[2] Antipsychotics • Amisulpride[3] • Aripiprazole[3] • Olanzapine[3] • Quetiapine[4] • Risperidone[3] • Ziprasidone[2] **Physical interventions:** • Acupuncture • ECT • Exercise • Yoga • Light therapy (for depression, not SAD)
Comparison	• Other active intervention (must also meet inclusion criteria above) • Treatment as usual • Waitlist • No treatment • Placebo
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 (usually defined as at least 50% improvement from the baseline score on a depression scale) • Relapse (number of participants who relapsed) The following depression scales will be included in the following hierarchy: • MADRS • HAMD • QIDS • PHQ • CGI (for dichotomous outcomes only) • CES-D • BDI

	• HADS-D (depression subscale) • HADS (full scale) **Acceptability/tolerability** • Discontinuation due to side effects (for pharmacological trials) • Discontinuation due to any reason (including side effects) **Important, but not critical, outcomes**: **Quality of life** • Quality of life (as assessed with a validated scale, including the 12-item/36-item Short-Form Survey [SF-12/SF-36], 26-item short version of the World Health Organization Quality of Life assessment [WHOQOL-BREF], EuroQoL [EQ5D], Quality of Life Depression Scale [QLDS], Quality of Life Enjoyment and Satisfaction Questionnaire [Q-LES-Q], Quality of Life Inventory [QoLI], and World Health Organization 5-item Well-Being Index [WHO-5]) **Personal, social, and occupational functioning** • Global functioning (as assessed with a validated scale, including Global Assessment of Functioning [GAF], Global Assessment Scale [GAS], and Social and Occupational Functioning Assessment Scale [SOFAS]) • Functional impairment (as assessed with a validated scale, including Sheehan Disability Scale [SDS], Social Adjustment Scale [SAS], and Work and Social Adjustment Scale [WSAS]) • Sleeping difficulties (as assessed with a validated scale, including Insomnia Severity Index [ISI] and Pittsburgh Sleep Quality Index [PSQI]) • Employment (for instance, % unemployed) • Interpersonal problems (as assessed with a validated scale, including Inventory of Interpersonal Problems [IIP]) Outcomes will be assessed at endpoint and follow-up (data for all available follow-up periods of at least 1-month post-intervention will be extracted and will be grouped into categories for analysis, for instance, 1-3 months, 4-6 months, 7-9 months, 10-12 months, 13-18 months, 19-24 months, and >2 years).
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried

	forward. Studies published between 2016 and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm Studies with <50% completion data (drop out of >50%) will be excluded.
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** A meta-analysis using a random-effects model will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical

	benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.
Heterogeneity (sensitivity analysis and subgroups)	No planned sub-group analysis
Notes.	Studies investigating further-line treatment of chronic depression will be considered under RQ 2.4 and any differences in efficacy due to chronic depression will be examined through sub-analysis in that review. *[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*

Treatment of depression: RQ 2.7 (depression with coexisting personality disorder)

Topic	First-line treatment or relapse prevention of depression with coexisting personality disorder
Review question	RQ. 2.7 For adults with depression and a coexisting personality disorder what are the relative benefits and harms of first-line treatment or relapse prevention with psychological, psychosocial, pharmacological and physical interventions alone or in combination?
Objectives	To identify the most effective first-line treatment or relapse prevention strategy for adults with depression and a coexisting personality disorder
Population	• Adults with depression and a coexisting personality disorder If some, but not all, of a study's participants are eligible for the review, then we will include a study if at least 80% of its participants are eligible for this review
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes) • Trials of further-line treatment following no/inadequate/limited response
Intervention	Interventions listed below are examples of interventions which may be included either alone or in combination. **Psychological interventions** • Behavioural therapies (including behavioural activation, behavioural therapy [Lewinsohn 1976], coping with depression group) • Cognitive and cognitive behavioural therapies (including CBT individual or group, problem solving, rational emotive behaviour therapy [REBT] and third-wave cognitive therapies individual or group) • Counselling (including emotion-focused therapy [EFT], non-directive/supportive/ person-centred counselling and relational client-centred therapy) • Family interventions/couples therapy • Psychodynamic psychotherapies (including short-term psychodynamic psychotherapy, long-term psychodynamic psychotherapy and psychodynamic counselling)

	• Self-help with or without support (including cognitive bibliotherapy with or without support, computerised CBT [CCBT] with or without support, computerised psychodynamic therapy with or without support) • Art therapy • Music therapy • Eye movement desensitization and reprocessing (EMDR) (for depression, not PTSD) **Psychosocial interventions** • Peer support (including befriending, mentoring, and community navigators) • Mindfulness, meditation or relaxation (including mindfulness-based stress reduction [MBSR]) **Pharmacological interventions** • Selective serotonin reuptake inhibitors • Tricyclic antidepressants • Serotonin-norepinephrine reuptake inhibitors • Other antidepressant drugs (including mirtazapine and trazodone) • Antipsychotics • Lithium • Omega-3 fatty acids **Physical interventions** • Acupuncture • ECT • Exercise • Yoga • Light therapy (for depression, not SAD)
Comparison	• Treatment as usual • Waitlist • No treatment • Placebo • Other active intervention (must also meet inclusion criteria above)
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 (usually defined as at least 50% improvement from the baseline score on a depression scale) • Relapse (number of participants who relapsed)

	The following depression scales will be included in the following hierarchy: • MADRS • HAMD • QIDS • PHQ • CGI (for dichotomous outcomes only) • CES-D • BDI • HADS-D (depression subscale) • HADS (full scale) **Acceptability/tolerability** • Discontinuation due to side effects (for pharmacological trials) • Discontinuation due to any reason (including side effects) **Important, but not critical, outcomes**: **Quality of life** • Quality of life (as assessed with a validated scale, including the 12-item/36-item Short-Form Survey [SF-12/SF-36], 26-item short version of the World Health Organization Quality of Life assessment [WHOQOL-BREF], EuroQoL [EQ5D], Quality of Life Depression Scale [QLDS], Quality of Life Enjoyment and Satisfaction Questionnaire [Q-LES-Q], Quality of Life Inventory [QoLI], and World Health Organization 5-item Well-Being Index [WHO-5]) **Personal, social, and occupational functioning** • Global functioning (as assessed with a validated scale, including Global Assessment of Functioning [GAF], Global Assessment Scale [GAS], and Social and Occupational Functioning Assessment Scale [SOFAS]) • Functional impairment (as assessed with a validated scale, including Sheehan Disability Scale [SDS], Social Adjustment Scale [SAS], and Work and Social Adjustment Scale [WSAS]) • Sleeping difficulties (as assessed with a validated scale, including Insomnia Severity Index [ISI] and Pittsburgh Sleep Quality Index [PSQI]) • Employment (for instance, % unemployed) • Interpersonal problems (as assessed with a validated scale, including Inventory of Interpersonal Problems [IIP]) Outcomes will be assessed at endpoint and follow-up (data for all available follow-up periods of at least 1-month post-intervention will be extracted and will be grouped into categories

	for analysis, for instance, 1-3 months, 4-6 months, 7-9 months, 10-12 months, 13-18 months, 19-24 months, and >2 years).
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from previous searches (pre-2016) will be carried forward. No restriction on date for the updated search, studies published between database inception and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm Studies with <50% completion data (drop out of >50%) will be excluded.
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** A meta-analysis using a random-effects model will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and

	completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.
Heterogeneity (sensitivity analysis and subgroups)	No sub-analyses are planned
Notes	Studies investigating further-line treatment of depression with coexisting personality disorder will be considered under RQ 2.4 and any differences in efficacy due to coexisting personality disorder will be examined through sub-analysis in that review

Treatment of depression: RQ 2.8 (psychotic depression)

Topic	Treatment of psychotic depression
Review question	RQ. 2.8 For adults with psychotic depression what are the relative benefits and harms of psychological, psychosocial, pharmacological and physical interventions alone or in combination (as first-line treatment or relapse prevention)?
Objectives	To identify the most effective first-line treatment or relapse prevention strategy for adults with psychotic depression
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) If some, but not all, of a study's participants are eligible for the review, then we will include a study if at least 80% of its participants are eligible for this review.
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes) • Depression occurring in a primary psychotic illness, such as schizophrenia or dementia • Trials of further-line treatment following no/inadequate/limited response
Intervention	Interventions listed below are examples of interventions which may be included either alone or in combination. **Psychological interventions** • Behavioural therapies (including behavioural activation, behavioural therapy [Lewinsohn 1976], coping with depression group) • Cognitive and cognitive behavioural therapies (including CBT individual or group, problem solving, rational emotive behaviour therapy [REBT] and third-wave cognitive therapies) • • Counselling (including emotion-focused therapy [EFT], non-directive/supportive/ person-centred counselling and relational client-centred therapy) • Family interventions/couples therapy

	• Interpersonal psychotherapy • Psychodynamic psychotherapies (including short-term psychodynamic psychotherapy, long-term psychodynamic psychotherapy and psychodynamic counselling) • Psychoeducational interventions (including psychoeducational group programmes) • Self-help with or without support (including cognitive bibliotherapy with or without support, computerised CBT [CCBT] with or without support, computerised psychodynamic therapy with or without support) • Art therapy • Music therapy • Eye movement desensitization and reprocessing (EMDR) (for depression, not PTSD) **Psychosocial interventions** • Peer support (including befriending, mentoring, and community navigators) • Mindfulness, meditation or relaxation (including mindfulness-based stress reduction [MBSR]) **Pharmacological interventions** • Selective serotonin reuptake inhibitors • Tricyclic antidepressants • Serotonin-norepinephrine reuptake inhibitors • Antipsychotics • Lithium • Omega-3 fatty acids **Physical interventions** • Acupuncture • ECT • Exercise • Yoga • Light therapy (for depression, not SAD)
Comparison	• Treatment as usual • Waitlist • No treatment • Placebo • Any other active comparison
Outcomes	**Critical outcomes**: **Efficacy** • Depression symptomology (mean endpoint score or change in depression score from baseline) • Response (usually defined as at least 50% improvement from the baseline score on a depression scale)

	• Remission (usually defined as a score below clinical threshold on a depression scale) • Relapse (number of people who returned to a depressive episode whilst in remission) The following depression scales will be included in the following hierarchy: • MADRS • HAMD • QIDS • PHQ • CGI (for dichotomous outcomes only) • CES-D • BDI • HADS-D (depression subscale) • HADS (full scale) **Acceptability/tolerability** • Discontinuation due to side effects (for pharmacological trials) • Discontinuation due to any reason (including side effects) **Important, but not critical, outcomes**: **Quality of life** • Quality of life (as assessed with a validated scale, including the 12-item/36-item Short-Form Survey [SF-12/SF-36], 26-item short version of the World Health Organization Quality of Life assessment [WHOQOL-BREF], EuroQoL [EQ5D], Quality of Life Depression Scale [QLDS], Quality of Life Enjoyment and Satisfaction Questionnaire [Q-LES-Q], Quality of Life Inventory [QoLI], and World Health Organization 5-item Well-Being Index [WHO-5]) **Personal, social, and occupational functioning** • Global functioning (as assessed with a validated scale, including Global Assessment of Functioning [GAF], Global Assessment Scale [GAS], and Social and Occupational Functioning Assessment Scale [SOFAS]) • Functional impairment (as assessed with a validated scale, including Sheehan Disability Scale [SDS], Social Adjustment Scale [SAS], and Work and Social Adjustment Scale [WSAS]) • Sleeping difficulties (as assessed with a validated scale, including Insomnia Severity Index [ISI] and Pittsburgh Sleep Quality Index [PSQI]) • Employment (for instance, % unemployed)

	• Interpersonal problems (as assessed with a validated scale, including Inventory of Interpersonal Problems [IIP]) Outcomes will be assessed at endpoint and follow-up (data for all available follow-up periods of at least 1-month post-intervention will be extracted and will be grouped into categories for analysis, for instance, 1-3 months, 4-6 months, 7-9 months, 10-12 months, 13-18 months, 19-24 months, and >2 years).
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the previous guideline)
Restriction by date?	All relevant studies from existing reviews from the 2009 guideline and from previous searches (pre-2016) will be carried forward. Studies published between 2016 and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm Studies with <50% completion data (drop out of >50%) will be excluded.
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** A meta-analysis using a random-effects model will be conducted to combine results from similar studies. An intention to treat (ITT) approach will be taken where possible.

	Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.
Heterogeneity (sensitivity analysis and subgroups)	No sub-analyses are planned
Notes	Studies investigating further-line treatment of psychotic depression will be considered under RQ 2.4 and any differences in efficacy due to psychotic depression will be examined through sub-analysis.

Access: RQ 3

Topic	Access to services for particular vulnerable groups
Review question	RQ.3 For adults (18 years and older) at risk of depression (or anxiety disorders) from particular vulnerable groups (older people, BME groups, LGBT 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?
Objectives	To identify the most effective service developments and interventions which are specifically designed to promote access
Population	Adults (18 years and older) identified as at risk of depression (or anxiety disorders*) from the following vulnerable groups - Older adults (mean age of 60 years or older) - BME groups - LGBT 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 Disorders NICE guideline will be undertaken. If some, but not all, of a study's participants are eligible for the review, then we will include a study if at least 80% of its participants are eligible for the review.
Exclude	• Trials of people with depression where the population does not fall into one of the particular vulnerable groups that are the focus of this review (older people, BME groups, LGBT groups and men) • Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Intervention	• 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
Critical outcomes	**Critical outcomes**: • Proportion of people from the target group who access treatment • Uptake of treatment **Important but not critical outcomes:** • Satisfaction, preference • Anxiety about treatment
Study design	• RCTs • Systematic reviews of RCTs
Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included unless the data can be extracted from elsewhere (for instance, from the CMHD guideline)
Restriction by date	All relevant studies from existing reviews from the Common Mental Health Disorders guideline and from previous searches (pre-2016) will be carried forward. No restriction on date for the updated search, studies published between database inception and the date the searches are run will be sought.
Minimum sample size	N = 10 in each arm
Study setting	Primary, secondary, tertiary and social care settings. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
The review strategy	**Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction will be double-coded. Discrepancies or difficulties with coding will be resolved through discussion between reviewers or the opinion of a third reviewer will be sought. **Data Analysis** A meta-analysis using a random-effects model will be conducted to combine results from similar studies.

	An intention to treat (ITT) approach will be taken where possible. Risk of bias will be assessed at the study level using the Cochrane risk of bias tool. This assessment includes: adequacy of randomisation (sufficient description of randomisation method, allocation concealment and any baseline difference between groups); blinding (of participants, intervention administrators and outcome assessors); attrition ('at risk of attrition bias' defined as a dropout of more than 20% and completer analysis used, or a difference of >20% between the groups); selective reporting bias (is the protocol registered, are all outcomes reported); other bias (for instance, conflict of interest in funding). Risk of bias will also be assessed at the outcome level using GRADE. For heterogeneity, outcomes will be downgraded once if I^2>50%, twice if I^2 >80%. For imprecision, outcomes will be downgraded using rules of thumb. If the 95% CI is imprecise i.e. crosses the line of no effect and the threshold for clinical benefit/harm, 0.8 or 1.25 (dichotomous) or -0.5 or 0.5 SMD (for continuous), the outcome will be downgraded. Outcomes will be downgraded one or two levels depending on how many lines it crosses. If the 95% CI is not imprecise, we will consider whether the criterion for Optimal Information Size is met (for dichotomous outcomes, 300 events; for continuous outcomes, 400 participants), if not we will downgrade one level.
Heterogeneity (sensitivity analysis and subgroups)	Where possible, the influence of the following subgroups will be considered: • Different subgroups within the LGBT category • Different subgroups within the BME category

Patient choice: RQ 4.0 (new question)

Topic	Patient choice
Review question	RQ. 4.0 What are the facilitators and barriers that can enhance or inhibit choice of treatment for adults with depression?
Objectives	To review the facilitators and barriers to patient choice in terms of treatment from the perspective of adults with depression and practitioners
Condition or domain being studied	• Adults with a diagnosis of depression according to DSM, ICD or similar criteria, or depressive symptoms as indicated by baseline depression scores on validated scales (and including those with subthreshold [just below threshold] depressive symptoms) If some, but not all, of a study's participants are eligible for the review, where possible data will be extracted for only eligible participants. If this is not possible then the study will be included if at least 80% of its participants are eligible for this review.
Exclude	• Trials of women with antenatal or postnatal depression • Trials of children and young people (mean age under 18 years) • Trials of people with learning disabilities • Trials of people with bipolar disorder • Trials of adults in contact with the criminal justice system (not solely as a result of being a witness or victim) • Trials that specifically recruit participants with a physical health condition in addition to depression (e.g. depression in people with diabetes)
Perspective	Service users (adults with depression) and practitioners
Phenomenon of interest	• Elements that adults with depression think are important to choice of pharmacological treatment • Elements that adults with depression think are important to choice of non-pharmacological treatment • Elements that adults with depression think are important to choice between pharmacological and non-pharmacological treatment • Factors or attributes (at the individual-, practitioner-, commissioner- or service- level) that can enhance or inhibit patient choice of treatment
Comparison	None
Study design	• Primary qualitative studies • Systematic reviews of primary qualitative studies (for identification of studies) Excluded: Commentaries, editorials, vignettes, books, policy and guidance, and non-empirical research

Include unpublished data?	Conference abstracts, dissertations and unpublished data will not be included
Restriction by date	Studies published between 2000 and the date the searches are run will be sought
Study setting	Primary, secondary, tertiary and social care settings. Studies from any OECD member country will be included. However, applicability to the UK service setting will be considered during data analysis and synthesis. Non-English-language papers will be excluded (unless data can be obtained from an existing review).
Evaluation	Experience and views of facilitators and barriers that can enhance or inhibit choice of treatment for adults with depression
The review strategy	**Data Extraction (selection and coding)** Citations from each search will be downloaded into EndNote and duplicates removed. Titles and abstracts of identified studies will be screened by two reviewers for inclusion against criteria, until a good inter-rater reliability has been observed (percentage agreement =>90%). Initially 10% of references will be double-screened. If inter-rater agreement is good then the remaining references will be screened by one reviewer. All primary-level studies included after the first scan of citations will be acquired in full and re-evaluated for eligibility at the time they are being entered into a study database (standardised template created in Microsoft Excel). **Data Synthesis** Qualitative data extraction and synthesis will be guided by a thematic analysis approach. This approach was selected as the review question is explorative in nature. Primary participant quotes pertaining to experience of choice of treatment will be extracted from the papers. Included studies will be divided between at least two reviewers, and each reviewer will examine the quotes in detail and develop their own coding framework. These individual analyses will be shared and a joint coding framework will be agreed and applied to the data. Quality at the individual study level will be assessed using the Critical Appraisal Skills Programme (CASP) quality-assessment tool, and each qualitative review finding will be assessed using the GRADE-CERQual (Confidence in the Evidence from Reviews of Qualitative Research) approach.
Notes	This is a new question added to the 2019 update scope