

# BMJ Open Comparative efficacy and acceptability of antidepressants in the long-term treatment of major depression: protocol for a systematic review and network meta-analysis

Kiyomi Shinohara,<sup>1</sup> Orestis Efthimiou,<sup>2</sup> Edoardo G Ostinelli,<sup>3</sup> Anneka Tomlinson,<sup>4</sup> John R Geddes,<sup>4,5</sup> Andrew A Nierenberg,<sup>6</sup> Henricus G Ruhe,<sup>7</sup> Toshi A Furukawa,<sup>1</sup> Andrea Cipriani<sup>4,5</sup>

**To cite:** Shinohara K, Efthimiou O, Ostinelli EG, *et al*. Comparative efficacy and acceptability of antidepressants in the long-term treatment of major depression: protocol for a systematic review and network meta-analysis. *BMJ Open* 2019;9:e027574. doi:10.1136/bmjopen-2018-027574

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-027574>).

Received 29 October 2018  
Revised 22 February 2019  
Accepted 27 March 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Kiyomi Shinohara;  
kiyomi.wb3@gmail.com

## ABSTRACT

**Introduction** Pharmacotherapy plays an important role in the treatment of major depression. At the initiation of antidepressant treatment, both improvement of symptoms in the short term and relapse prevention in the long term should be taken into account. However, there is insufficient evidence regarding the efficacy and the acceptability of continuation/maintenance treatments and the relative efficacy/acceptability of antidepressants.

**Objective** We will conduct a pairwise meta-analysis and a network meta-analysis (NMA) to examine the relative efficacy, tolerability and acceptability of antidepressants in the long-term treatment of major depression.

**Methods and analysis** We will include double-blind randomised controlled trials comparing any of the following antidepressants, which we included in our previous NMA of the acute treatment for major depression, with placebo or with another active drug for long-term treatment of major depression: agomelatine, amitriptyline, bupropion, citalopram, clomipramine, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilnacipran, milnacipran, mirtazapine, nefazodone, paroxetine, reboxetine, sertraline, trazodone, venlafaxine, vilazodone and vortioxetine. Our primary outcomes will be sustained response and all-cause dropouts. We will include four types of designs that are used to investigate long-term treatment. We will conduct two main analyses. First, we will conduct a pairwise meta-analysis comparing all antidepressants versus placebo to investigate whether continuing antidepressants after achieving a positive response in the acute-phase treatment is beneficial and/or safe. Second, we will conduct an NMA to examine the comparative efficacy and acceptability of the drugs. We will use a novel approach that will combine the results of acute-phase treatment NMA with long-term treatment studies to include all related designs in the NMA. We will ensure the validity of combining different designs and our new approach by checking the distribution of important effect modifiers and consistency of network.

**Ethics and dissemination** This study did not require ethical approval. We will disseminate our findings by publishing results in a peer-reviewed journal.

## Strengths and limitations of this study

- Our review will answer one of the most important questions about the treatment of depression: which drug, if any, should be preferred in terms of efficacy, tolerability and acceptability during the acute phase and the continuation/maintenance treatment phase for major depression?
- We will use a novel approach to combine the results of acute-phase treatment network meta-analysis (NMA) with long-term treatment studies, thus enabling the inclusion of all related designs in the NMA.
- Our primary outcomes, 'sustained response' and 'all-cause dropouts', may have slightly different meanings, depending on the study design.
- Studies examining some tricyclics for relapse prevention will be excluded from our NMA because our focus is on the 21 antidepressants previously included in the NMA for the acute treatment of depression.

**PROSPERO registration number** CRD42018114561; Pre-results.

## INTRODUCTION

Depression is estimated to affect over 300 million people across the world and contributes 7.5% of all years lived with disability in the overall global burden of disease.<sup>1</sup> It runs a chronic-recurrent course as 30%–40% of cases experience recurrent episodes within a year.<sup>2</sup> Studies show that the years of life lived with disability due to depressive disorders increased by 37.5% between 1990 and 2010. Two-thirds of this increase comes from population growth, and one-third came from population ageing without a change in age-specific prevalence,

as more people live to the age where they can experience depressive disorders.<sup>3,4</sup>

Pharmacotherapy plays an important role in the treatment of depression. At the initiation of treatment, both improvement of symptoms in the short term and relapse prevention in the long term should be taken into account. A comprehensive systematic review of 522 double-blind randomised controlled trials (RCTs) has recently demonstrated that antidepressants are efficacious for the acute symptoms of depression and that some antidepressants may be more efficacious or more acceptable than others.<sup>5</sup> However, there is insufficient evidence regarding long-term treatment. Although antidepressants also reduce relapse rates,<sup>6–8</sup> there is no universal consensus on how long continuation/maintenance treatment should last.<sup>9–13</sup> There is also no consensus on which antidepressant, if any, should be preferred as a first option when envisaging the long-term use.<sup>14</sup> The lack of evidence raises concerns about the benefits and harms of the long-term use of antidepressants.<sup>15</sup> It is therefore currently difficult to choose an antidepressant in the acute phase while taking into account long-term efficacy and acceptability.

In this study, we will conduct a pairwise meta-analysis and a network meta-analysis (NMA) to examine the relative efficacy, tolerability and acceptability of antidepressants in the long-term treatment of major depression. First, using the pairwise meta-analysis, we will obtain estimates of efficacy, tolerability and acceptability of continuing antidepressants, in comparison with discontinuing the antidepressant after achieving a positive response in the acute phase of treatment. Second, using NMA, we will

estimate the relative efficacy, tolerability and acceptability of various antidepressants, in both the acute phase and the continuation/maintenance phase of treatments.

## METHODS

We used the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols (PRISMA-P).<sup>16</sup>

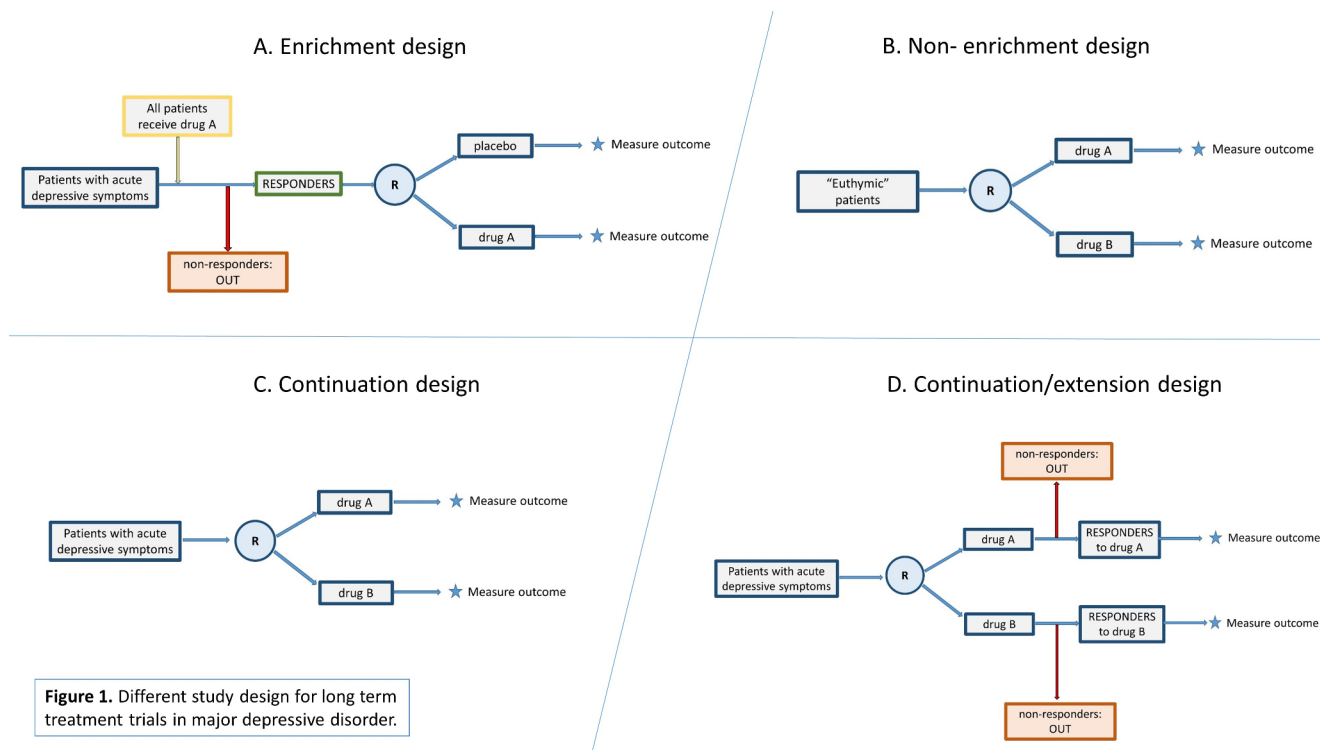
### Eligibility criteria

#### Types of studies

We will include double-blind RCTs, in which any of the 21 antidepressants described further (see the Types of interventions section) was compared as monotherapy with placebo or with another active drug for the long-term treatment of unipolar major depression.

We define long-term treatment as the duration of treatment of more than 12 weeks.<sup>17</sup> We conceptualise the phases of treatment for major depression as consisting of the acute-phase treatment lasting up to 3 months, and the subsequent continuation/maintenance phase.<sup>18</sup> The distinction between continuation and maintenance is more conceptual than pragmatic,<sup>6</sup> and we will include both of them in the current analyses.

In the scientific literature, there are four types of methodological designs used to assess the effectiveness of drugs used as long-term treatment (figure 1). The most commonly used is an 'enrichment design'.<sup>19</sup> In this type of study (type A in figure 1), patients who responded to an open-label acute treatment (*run-in phase*) are subsequently randomised to continue the active drug or to switch to placebo (ie, the randomised sample is enriched



**Figure 1.** Different study design for long term treatment trials in major depressive disorder.

**Figure 1** Different study designs for long-term treatment trials in a major depressive disorder.

with people who have responded acutely to the active agent). The second study design (type B) is a 'non-enrichment design', where euthymic patients (ie, in remission) are included and randomised (in this type of study, there is no run-in or open-label period). The third design (type C) is 'continuation design', in which patients with acute depressive symptoms are randomly allocated to receive treatment or placebo (or another active treatment) and then followed up for a long-term period (ie, more than 3 months). Another type of continuation design is the so-called 'continuation/extension design' (type D), which includes an extension phase, where only participants who responded to the acute treatment they have been allocated to (drug and placebo/another active treatment) are followed up in a double-blind fashion. Quasi-randomised or cross-over trials will be excluded, while cluster RCTs will be included when the clustering effect can be taken account of.

### Types of participants

We will include patients aged 18 years or older, of both sexes with unipolar major depression diagnosed by using any standard operationalised criteria, such as Feighner criteria, Research Diagnostic Criteria, *Diagnostic and Statistical Manual of Mental Disorders* (DSM)-III, DSM-III-R, DSM-IV, DSM-5 and International Classification of Diseases, 10th Revision (ICD-10). Studies in which 20% or more of the participants suffering from bipolar disorder, treatment resistant depression, peripartum depression or psychotic depression will be excluded. We will include chronic depression (depressive episode lasting longer than 2 years) and 'double depression' (major depression superimposed on dysthymic disorder) as long as the patients met the diagnostic criteria of major depression at the commencement of acute treatment. We will exclude RCTs that have included patients with a concurrent primary diagnosis of another Axis I or II disorder. However, a concurrent secondary diagnosis of another psychiatric disorder will not be considered as exclusion criteria. As depressive symptoms can be influenced by physical illness,<sup>20</sup> RCTs focusing on patients with a concomitant medical illness will be excluded. We will also exclude seasonal affective disorders as their seasonal recurrent pattern differs from general depression.<sup>21</sup> We will include studies that allow the use of rescue medications, if these medications were made equally available to all treatment groups.

### Types of interventions

We will include the 21 antidepressants, which showed greater efficacy than placebo in the acute treatment of major depression in our previous review.<sup>5</sup> In this last, all antidepressants currently marketed in the USA, Europe or Japan were included: agomelatine, bupropion, citalopram, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilnacipran, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, venlafaxine, vilazodone and vortioxetine. In addition to these,

four drugs were selected: two tricyclic antidepressants (amitriptyline and clomipramine) listed in the WHO Model List of Essential Medicines, and trazodone and nefazodone, as their effect and tolerability profiles differ from others.<sup>5</sup> We suggest that it is reasonable to focus on these 21 antidepressants because, in the long-term treatment of depression, a continuation of the drug to which the patient responded in the acute treatment phase is recommended. We chose to focus on these 21 antidepressants that have evidence of efficacy for acute treatment rather than evaluate a variety of drugs with unclear evidence for acute treatment.

We will obtain information about the interventions of interest both from head-to-head and placebo controlled trials. We will include studies that used the eligible drugs as monotherapy for acute treatment in enrichment and non-enrichment designs. Studies that used non-pharmacological interventions (eg, electroconvulsive therapy and psychotherapy) or ineligible drugs for the acute treatment in more than 20% of patients will be excluded.

### Outcome measures

Primary outcomes:

- ▶ *Sustained response (efficacy)*, measured as the proportion of patients who had responded in the acute-phase treatment and who subsequently did not have a depressive relapse during the continuation/maintenance phase, up to the end of the study duration. We will use the original authors' definitions of 'response' or 'relapse/recurrence'.
- ▶ *All-cause dropouts (acceptability)*, measured as the proportion of patients who left the trial early due to any reason up to the end of the study duration.

Secondary outcomes:

- ▶ *Dropouts due to adverse events (tolerability)*, measured as the proportion of patients who left the trial early, either during the acute-phase treatment or during the continuation/maintenance treatment, due to any adverse events (including specific adverse events and withdrawal symptoms).
- ▶ *Occurrence of specific adverse events*, as reported in the original studies.
- ▶ *Suicidality*, measured as the number of patients who deliberately self-harmed, attempted or completed suicide.

### Search strategy and data management

#### Search strategy

We will identify all published, unpublished and ongoing RCTs that compared the efficacy and/or acceptability of one drug with another or with placebo in the treatment of major depression. We will use search terms as follows: depress\* or dysthymi\* or adjustment disorder\* or mood disorder\* or "affective disorder" or "affective symptoms" (a full search strategy is in the online supplementary file). The following sources will be searched: Cochrane Central Register of Controlled Trials, AMED, CINAHL, Embase, LILACS database, MEDLINE, MEDLINE In-Progress,



PsycINFO and the UK National Research Register. We will also hand-search a list of trial databases from drug-approving agencies (Food and Drug Administration (USA), Medicines and Healthcare Products Regulatory Agency (UK), European Medicines Agency (Europe), Medicines Evaluation Board (the Netherlands), Medical Products Agency (Sweden), Pharmaceuticals and Medical Devices Agency (Japan) and Therapeutic Goods Administration (Australia)) and international trial registries (clinicaltrials.gov, ISRCTN, National Research Register in the UK, the Netherlands Trial Register, EUDRACT, UMIN-CTR, JapicCTI and JMACCT, Australian Clinical Trials Registry and WHO International Clinical Trials Registry Platform). We will contact the National Institute for Clinical Excellence (UK), the Institut für Qualität und Wirtschaftlichkeit in Gesundheitswesen (Germany) and representatives of pharmaceutical companies to obtain unpublished information (we will also check their websites). No date or language restrictions will be applied to any of the searches. All relevant authors will be contacted to supplement any unpublished studies or incomplete reporting in the original papers and will be reminded twice.

We will include Chinese trials as long as they were registered in an international database regardless of their language. However, we will exclude them when they were registered only in a Chinese database. Although increasing numbers of clinical trials have been reported from mainland China,<sup>22</sup> there is a concern about their quality.<sup>23 24</sup> It has been under debate whether Chinese databases should be searched when conducting a systematic review.<sup>25</sup> Moreover, recent studies have shown that the trials indexed in Chinese databases generally suffer from numerous quality issues,<sup>24</sup> and evaluating their credibility and integrity is beyond the scope of this work.<sup>25</sup> We therefore decided to exclude such studies to avoid any potential bias that might arise by their inclusion.

### Study selection

Two reviewers will independently review references and abstracts retrieved by the search. If both reviewers agree that a trial does not meet the eligibility criteria, it will be excluded. We will obtain the full text of all remaining articles and use the same eligibility criteria to determine the final selection. Any disagreements will be resolved via discussion with a third member of the review team.

### Data extraction

Two reviewers will independently extract data and evaluate the risk of bias for each selected trial. We will use a structured data extraction template designed for this study and check the reliability of the data extraction. Information extracted will include trial characteristics (eg, lead author, journal, publication year, design, blinding and inclusion criteria), participant characteristics (eg, diagnostic criteria for depression, age range, setting and average number of previous episodes), intervention details (eg, duration of treatment, dose ranges and mean doses of study drugs) and outcome measures.

In enrichment trials, we will also extract the details of acute treatment before randomisation (ie, length of treatment prior to randomisation and dosage of drug), the method of tapering off the active drug after randomisation in the placebo arm (abrupt or gradual) and the definition of remission/response required for entry.

### Length of trial

We will include any trial with a duration greater than 3 months. For the main analysis, we will consider the longest endpoint data. We will classify studies into several groups based on the length of treatment (from the initiation of acute treatment to the end of continuation treatment: >3 and ≤6 months, >6 and ≤12 months, >12 and ≤24 months, and >24 months) to assess, using subgroup analyses, whether the duration of treatment leads to any differences in benefit.<sup>6</sup>

### Comparability of dosages

We will include both fixed and flexible dose designs, and only include arms randomising patients to drugs within the licensed dose range. We will consider all comparisons within licensed ranges to be fair comparisons.

### Risk of bias assessment

We will assess risk of bias of the included studies with regard to the primary outcome, using the tool described in the *Cochrane Collaboration Handbook* as a reference guide.<sup>26</sup> The assessment will be performed by two independent raters. If the raters disagree, the final rating will be made by consensus with the involvement of another member of the review group. The following domains will be assessed: sequence generation, allocation concealment, blinding of therapists and participants, blinding of outcome assessors, incomplete outcome data, selective outcome reporting and other source of bias (eg, sponsorship bias). We will contact the authors to obtain missing information, if necessary. Overall risk of bias of each study will be categorised as follows<sup>5 27</sup>: studies will be classified as having overall low risk of bias if none of the domains above are rated as high risk of bias and three or less are rated as unclear risk; and overall moderate if one domain is rated as high risk of bias, or none are rated as high risk of bias but four or more are rated as unclear risk. All other studies will be assumed to pertain to the overall high risk of bias group.

### Statistical analysis

#### Synthesis of results

We will present the trial and study population characteristics across all eligible trials, describing the types of designs, comparisons and important variables, such as year of publication, age, sponsorship and clinical setting. We will analyse dichotomous outcomes on an intention-to-treat basis, assuming that all dropouts from the treatment had negative outcomes (ie, not sustaining response).

### Pairwise meta-analysis

As we aim to answer two clinical questions in this review, we will conduct two main analyses. The first clinical question relates to whether continuing antidepressants after achieving a positive response in the acute-phase treatment is beneficial and/or safe. To answer this question, we will conduct a pairwise meta-analysis comparing all antidepressants (ie, all drugs grouped together) versus placebo to investigate the effects of antidepressants in long-term treatment (analysis 1). For this analysis, we will only use enrichment-design studies (type A in figure 1). We will use a random effects meta-analysis model<sup>28</sup> to estimate odds ratios. We will estimate the heterogeneity parameter using restricted maximum likelihood and present the 95% CIs using the Knapp-Hartung method. In this analysis, we will consider all primary and secondary outcomes. We will use R to perform all analyses.<sup>29</sup> Should the event be rare for some of the outcomes (eg, zero events in one or both treatment arms for some studies), we will consider alternative methods for meta-analysis.<sup>30</sup>

### Network meta-analysis

The second clinical question is about how the various antidepressants compare with each other, when both the acute and the continuation/maintenance phases of the treatment are taken into account. For this (analysis 2), we will conduct NMAs to examine the comparative efficacy and acceptability of the drugs. NMA is a method for jointly synthesising direct and indirect evidence for each treatment comparison. It thus allows more studies to contribute towards each outcome, for each treatment comparison.<sup>31 32</sup> However, synthesising the results of the four different designs that we expect to find (figure 1) via standard NMA methods poses a problem regarding transitivity, which is the underpinning assumption of the NMA model.<sup>33</sup> According to this assumption, there are no important differences between studies in clinical and methodological aspects that might affect the outcome, other than the treatments being compared. In the present context, however, we cannot automatically assume that participants allocated to different drugs have similar characteristics across the different study designs. For instance, participants of an enrichment study will all be responders to a particular drug A; participants of another enrichment study will all be responders to a different drug (ie, drug B). To assume that the patients who acutely responded to drug A are identical to those that acutely responded to drug B may be wrong. In addition, such an assumption (which is equivalent to assuming that drugs A and B are equally or similarly effective) defeats the purpose of this review, which is to explore possible differences between the drugs. Thus, a standard NMA approach including all designs will not be appropriate.

To overcome this problem, we will start by performing a random effects NMA only using type B and C designs, if enough such studies are available (NMA 1). We will do this after assuming that treatment effects are similar in euthymic patients, and patients with acute depressive

symptoms. In our analysis, we will assume a common heterogeneity variance for all treatment comparisons in the network. We will perform this NMA for all primary and secondary outcomes.

We anticipate, however, that we can find fewer studies of types B and C (enrichment-design studies are much more frequent in the scientific literature), possibly employing only a subset of the 21 drugs. In this case, NMA one will not be able to answer the clinical question at hand. If this is the case, we will follow a 'chain-of-evidence' approach to synthesise the totality of the evidence.<sup>34</sup> Using this approach, we will amalgamate the results of the NMA of acute treatment<sup>5</sup> with type A, B, C and D studies to obtain an answer to our clinical question.

More concretely, the outcome we are interested in is sustained response, which corresponds to 'long term response AND acute response'. Thus, we are interested in the probability:

$P(\text{long-term response AND acute response, under treatment } T)$ , for all different treatments, for the general population of patients. Type B, C and D studies inform this outcome directly. However, type A studies only give information for the outcome 'long-term response GIVEN acute response', that is, they measure

$$P(\text{long-term response} \mid \text{acute response, under treatment } T).$$

These two probabilities are related via the formula:

$$\begin{aligned} &P(\text{long-term response AND acute response, under treatment } T) \\ &= P(\text{long-term response} \mid \text{acute response, under treatment } T) \\ &\times P(\text{acute response, under treatment } T) \end{aligned} \quad (1)$$

Thus, in order to allow type A studies (which we expect to be the majority of studies) to also inform the quantity of interest, that is, the left part of equation (1), we will need information about  $P(\text{acute response, under treatment } T)$ . This probability is not informed by any of the study designs in figure 1. We will thus inform this term using the results from the existing NMA on acute treatment.<sup>5</sup>

The steps of our chain-of-evidence analysis plan are as follows:

1. We will use the results of the NMA of acute treatment<sup>[5]</sup> to estimate  $P(\text{acute response, under treatment } T)$  for each treatment  $T$ .
2. We will use type A studies to estimate  $P(\text{long-term response} \mid \text{acute response, under treatment } T)$  for each treatment.
3. We will use the estimates from steps 1 and 2, together with equation (1), to estimate for each treatment  $T$  the probability  $P(\text{long-term response AND acute response, under treatment } T)$ .
4. We will use results from step 3 to estimate odds ratios for all treatments versus an (arbitrarily chosen) control treatment.
5. We will use type B, C and D studies to obtain study-specific estimates for the odds ratios of the corresponding treatments, for sustained response.

6. We will jointly synthesise the estimated effect sizes from steps 4 and 5 in another random-effects NMA (NMA 2), assuming a common heterogeneity for all comparisons in the network. We will report the odds ratios for all treatment comparisons, along with their corresponding uncertainty.

Note that during steps 1, 2 and 3 of this approach, given the information provided by type A studies, we will also be able to estimate the effects of composite treatments, for example, drug A for the acute phase and placebo for the continuation. We will perform this analysis for the first primary outcome only, sustained response. All analyses will be performed in R.

#### Assessment of heterogeneity (pairwise meta-analysis)

In the pairwise meta-analysis, we will check the possibility of heterogeneity by visually inspecting the forest plots. We will then report the estimated value for the heterogeneity variance and compare it with the corresponding empirical distribution.<sup>35,36</sup> Finally, we will report the  $I^2$  statistic,<sup>37</sup> which corresponds to the percentage of the variability due to heterogeneity rather than a sampling error. In the NMA, we will estimate the heterogeneity variance and judge its magnitude by comparing it with empirical distributions.<sup>35,36</sup>

#### Assessment of the transitivity assumption (NMA)

As discussed previously, transitivity is a fundamental assumption for NMA. The distribution of potential effect modifiers should be balanced across treatment comparisons.

For NMA 1, we will check for significant differences in the distribution of the following factors among different comparisons: severity of depression before randomisation, age, average number of previous episodes and clinical setting.

Checking transitivity for NMA 2 will require one additional step, namely, we will check whether important effect modifiers have similar distributions between the studies included in the NMA for acute treatment<sup>5</sup> and the enrichment studies (type A) before enrichment. This is needed to ensure that step 3 of the analysis is valid, that is, that the different terms of equation (1) are estimated in similar populations. Also, we will check whether the participants' characteristics and severity of depression are similar enough between the studies included in acute NMA and those with continuation designs (ie, type C and D).

#### Assessment of inconsistency

Another method for assessing the validity of a NMA is to check whether the network is consistent, that is, that there is statistical agreement between direct and indirect evidence on the same comparisons. Consistency is a reflection of the transitivity of the NMA, and a breach in the transitivity assumption may lead to important inconsistencies in the network.<sup>32</sup> In NMA 1, we will evaluate the consistency by using two approaches: via the back-calculation

method to assess local inconsistency,<sup>38</sup> and the design-by-treatment test to assess the global inconsistency in the network.<sup>39</sup> In case we identify important inconsistency, we will explore possible sources via subgroup and meta-regression analyses (described further). If the extent of inconsistency is deemed important, we will abstain from performing any further analyses. In addition, this analysis assumes that treatment effects are similar in euthymic patients and in patients with acute depressive symptoms. We will explore the plausibility of this assumption by checking whether the estimated treatment effects in type B and C studies are in statistical agreement.

In NMA 2, we will assess inconsistency in step 6 by checking for differences between the estimates obtained using type A studies in conjunction with the NMA of acute treatment<sup>5</sup> versus the estimates obtained from type B, C and D studies.

#### Assessment of publication bias and small study effect

In order to assess whether small studies report higher effect estimates than larger studies (due to publication bias/small study effect) in the pairwise meta-analysis, we will visually examine the contour-enhanced funnel plots and perform the Peters test.<sup>40</sup> For NMA 1, we will use comparison-adjusted funnel plots, after assuming that newer treatments are favoured.<sup>41</sup>

We will also evaluate the influence of publication bias by performing subgroup analyses of published versus unpublished data.

#### Sensitivity analyses and subgroup analyses

In order to investigate possible heterogeneity in the pairwise meta-analysis, we will examine the effect of the following study-level characteristics, in meta-regressions or subgroup analyses: (1) study year, (2) length of acute treatment at the time of randomisation, (3) length of remission required before randomisation, (4) average age of patients, (5) average number of previous episodes, (6) clinical setting and (7) type of discontinuation of the antidepressant.

For NMA 2, we will conduct a subgroup analysis after grouping trials (if possible) according to their durations (>3 and ≤6 months, >6 and ≤12 months, >12 and ≤24 months, and >24 and ≤36 months). If possible (ie, depending on the data we will obtain), we will instead perform a multiple-outcome NMA to jointly synthesise treatment effect estimates at multiple time points.<sup>42,43</sup> We will also perform an analysis aiming to assess novelty effects.<sup>44</sup>

Finally, depending on data availability, we will exclude continuation non-extension design (ie, type C studies) because they possibly included people who have not responded after the acute treatment phase, and they would have been excluded in enrichment and continuation-extension design.

Depending on other data that might become available, we will perform further post hoc analyses to assess the



influence of additional factors possibly related to treatment effects.

### GRADE quality assessment

The risk of bias of individual trials and the certainty of evidence contributing to network estimates of the primary outcomes will be assessed by Confidence in Network Meta-Analysis (Software) (CiNEMA) using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework<sup>45</sup> for NMA 1. The full article will be reported following standardised guidelines.<sup>46</sup>

### Patient and public involvement

There was no patient and public involvement in the design of this study.

### DISCUSSION

This review aims to provide the best evidence summary about the long-term treatment of major depression. By extending the current evidence regarding acute treatment, the findings of this review will hopefully give us a better understanding of the relative efficacy, acceptability and tolerability of antidepressants while taking into account both the acute and long-term treatments.

We will address two related clinical questions regarding continuation/maintenance treatment for major depression. Once remitted from the acute phase of antidepressant therapy, should one continue the antidepressant? If yes, in view of such long-term treatment, which antidepressants, if any, would be preferred at the time of the initiation of the acute-phase treatment? Although previous systematic reviews of long-term treatment of antidepressants showed the efficacy of continuation treatment, they focused on a particular design<sup>6 7 19 47</sup> or analysed various designs together without accounting for their differences.<sup>8</sup> Moreover, the relative efficacy of various drugs has never been studied.

However, synthesis of studies of the long-term treatment of major depression using different designs into the NMA framework poses a number of challenges. For example, there are two issues in pooling enrichment-design studies examining different drugs in the NMA. First, patients who have responded to drug A and have then been randomised to continue drug A or switch to placebo may be different in known and/or unknown characteristics from those who have responded to drug B and then randomised to B or placebo, that is, there is an issue regarding the joint randomisability of such patients,<sup>32 48</sup> which might render the transitivity assumption invalid. Second, even if we could somehow overcome the transitivity issue, the enrichment studies would only form a so-called 'star network', with the placebo being the only common comparator, that is, there will be no closed loops in the network, and all the comparisons among the active drugs will be indirect and hence will probably give imprecise estimates. Pooling together enrichment-design studies and continuation-design studies with or

without extension, would close the network but would pose yet another issue; the former will have randomised euthymic patients, while the latter will have randomised acutely symptomatic patients who responded to specific drugs, again posing threats to the joint randomisability assumption.

To overcome these issues, we will use a novel approach that will combine the results of acute-phase treatment NMA<sup>5</sup> with the enrichment-design studies, and then further pool them with studies of the other designs. However, this new methodology has its own weaknesses. First, our new approach of combining the results of the NMA of acute treatment and those of enrichment-design trials does not preserve randomisation among the arms because we will compare groups that were not originally randomised into the arms. We will use the probability of response in the previous NMA of acute treatment and combine it with evidence from enrichment-design studies to obtain the probability of sustaining response/remaining in treatment in long-term treatment. We propose this new approach because it will enable us to use all related designs to answer the clinically relevant questions. Second, we chose 'sustained response' and 'all-cause dropouts' as our primary outcomes because they are clinically important and are expected to be reported across various designs. However, these outcomes may have slightly different meanings, depending on the design. In the enrichment design, the 'sustaining response' group will include participants who have responded to the acute treatment and remain free of relapse up to the end of continuation treatment, while in continuation design, the sustained response group may include participants who met the criteria of response at the end of study. We will describe these differences, if there are any, and discuss their impact on the review findings. Finally, we decided to focus on the 21 antidepressants that we included in the previous NMA for the acute treatment of depression.<sup>5</sup> We are aware that several studies examining tricyclics (eg, nortriptyline and imipramine) for relapse prevention will then be excluded from our NMA. This will not, however, undermine the clinical importance of our study because tricyclics and the older antidepressants are not recommended as first-line treatment for depression.

Notwithstanding these weaknesses, our review will summarise the best available evidence to answer one of the most important questions about the treatment for depression: which drug, if any, should be preferred in view of efficacy, tolerability and acceptability of the agents through the acute-phase and continuation/maintenance treatments for major depression? We believe that the results of our NMA will help clinicians and patients make better-informed choices from a long-term perspective, when they initiate antidepressant treatments for their acute depressive symptoms.

**Author affiliations**

<sup>1</sup>Departments of Health Promotion and Human Behavior and of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan

<sup>2</sup>Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

<sup>3</sup>University of Milan, Milano, Lombardia, Italy

<sup>4</sup>Department of Psychiatry, University of Oxford, Oxford, UK

<sup>5</sup>Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK

<sup>6</sup>Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA

<sup>7</sup>Department of Psychiatry, Radboud University, Nijmegen, the Netherlands

**Contributors** KS, AT, AC, TAF and OE designed this study and drafted the protocol. AC developed the search strategies. KS, EGO, AT, TAF and AC independently screened potential studies and extracted data from the included studies, assessed the risk of bias and summarised the evidence. JRG, AAN and HGR arbitrated in cases of disagreement and ensured the absence of errors. OE performed the statistical analysis. All authors approved the publication of this protocol.

**Funding** This study was supported in part by JSPS Grant-in-Aid for Scientific Research (Grant Number 17K19808) to TAF. Andrea Cipriani is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017-08-ST2-006) and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005). The views expressed are those of the authors and are not necessarily those of the UK National Health Service, the NIHR or the UK Department of Health. The funders had no role in developing the protocol.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** We will disseminate our findings by publishing in a peer-reviewed journal.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

**REFERENCES**

- World Health Organization. Depression and Other Common Mental Disorders: Global Health Estimates. 2017. Licence: CC BY-NC-SA 3.0 IGO [http://www.who.int/mental\\_health/management/depression/prevalence\\_global\\_health\\_estimates/en/](http://www.who.int/mental_health/management/depression/prevalence_global_health_estimates/en/) (accessed 26 Oct 2018).
- Kessler RC, Bromet EJ. The epidemiology of depression across cultures. *Annu Rev Public Health* 2013;34:119–38.
- Ferrari AJ, Charlson FJ, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med* 2013;10:e1001547.
- World Health Organisation. *Depression: fact sheet*. Geneva, 2017. <http://www.who.int/mediacentre/factsheets/fs369/en/>. (Accessed 26 Oct 2018).
- Cipriani A, Furuwaka TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 2018;391:1357–66.
- Geddes JR, Carney SM, Davies C, et al. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *Lancet* 2003;361:653–61.
- Hansen R, Gaynes B, Thieda P, et al. Meta-analysis of major depressive disorder relapse and recurrence with second-generation antidepressants. *Psychiatr Serv* 2008;59:1121–30.
- Sim K, Lau WK, Sim J, et al. Prevention of relapse and recurrence in adults with major depressive disorder: Systematic review and meta-analyses of controlled trials. *Int J Neuropsychopharmacol* 2016;19:pyv076–13.
- National Institute for Health and Clinical Excellence. Depression: the treatment and management of depression in adults (update) (Clinical guideline 90). 2009 <https://www.nice.org.uk/guidance/CG90> (Accessed 26 Oct 2018).
- Clear A, Pariante CM, Young AH, et al. Evidence-based guidelines for treating depressive disorders with antidepressants: A revision of the 2008 British Association for Psychopharmacology guidelines. *J Psychopharmacol* 2015;29:459–525.
- Malhi GS, Bassett D, Boyce P, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. *Aust N Z J Psychiatry* 2015;49:1087–206.
- Kennedy SH, Lam RW, McIntyre RS, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 3. Pharmacological Treatments. *Can J Psychiatry* 2016;61:540–60.
- American Psychiatric Association. Practice Guideline for the Treatment of Patients with Major Depressive Disorder. 3rd ed, 2010. <https://psychiatryonline.org/guidelines>. (Accessed 26 Oct 2018).
- Furuwaka TA, Cipriani A, Barbui C, et al. Long-term treatment of depression with antidepressants: a systematic narrative review. *Can J Psychiatry* 2007;52:545–52.
- McCormack J, Korownyk C. Effectiveness of antidepressants. *BMJ* 2018;360:k1073.
- Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;350:g7647.
- Cipriani A, Hawton K, Stockton S, et al. Lithium in the prevention of suicide in mood disorders: updated systematic review and meta-analysis. *BMJ* 2013;346:f3646.
- Kupfer DJ. Long-term treatment of depression. *J Clin Psychiatry* 1991;52 Suppl:28–34.
- Deshauer D, Moher D, Fergusson D, et al. Selective serotonin reuptake inhibitors for unipolar depression: a systematic review of classic long-term randomized controlled trials. *CMAJ* 2008;178:1293–301.
- Šprah L, Dernovšek MZ, Wahlbeck K, et al. Psychiatric readmissions and their association with physical comorbidity: a systematic literature review. *BMC Psychiatry* 2017;17:1–17.
- Gartlehner G, Nussbaumer B, Gaynes BN, et al. Second-generation antidepressants for preventing seasonal affective disorder in adults. *Cochrane Database Syst Rev* 2015;11:CD011268.
- Wang M. Clinical trials and drug approvals continue to accelerate in China. *Lancet Oncol* 2017;18:855.
- Yao C. Clinical trial in China: The status and challenge of data management and statistical analysis. *J Evid Based Med* 2018;11:3–6.
- Tong Z, Li F, Ogawa Y, et al. Quality of randomized controlled trials of new generation antidepressants and antipsychotics identified in the China National Knowledge Infrastructure (CNKI): a literature and telephone interview study. *BMC Med Res Methodol* 2018;18:96.
- Cohen JF, Korevaar DA, Wang J, et al. Should we search Chinese biomedical databases when performing systematic reviews? *Syst Rev* 2015;4:23.
- Higgins JPT GS. Cochrane handbook for systematic reviews of interventions version 5.1.0. <https://training.cochrane.org/handbook> (Accessed 26 Oct 2018).
- Furuwaka TA, Salanti G, Atkinson LZ, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *BMJ Open* 2016;6:e010919.
- Nikolakopoulou A, Mavridis D, Salanti G. Demystifying fixed and random effects meta-analysis. *Evid Based Ment Health* 2014;17:53–7.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Found: Stat. Comput, 2014. <http://www.r-project.org>.
- Eftthimiou O. Practical guide to the meta-analysis of rare events. *Evid Based Ment Health* 2018;21:72–6.
- Riley RD, Jackson D, Salanti G, et al. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ* 2017;358:j3932.
- Eftthimiou O, Debray TP, van Valkenhoef G, et al. GetReal in network meta-analysis: a review of the methodology. *Res Synth Methods* 2016;7:236–63.
- Cipriani A, Higgins JP, Geddes JR, et al. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013;159:130–7.
- Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Stat Med* 2003;22:2995–3016.
- Turner RM, Jackson D, Wei Y, et al. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Stat Med* 2015;34:984–98.
- Rhodes KM, Turner RM, Higgins JP. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of



- continuous outcome data. *J Clin Epidemiol* 2015;68:52–60.
37. Higgins JP, Thompson SG, Deeks JJ, *et al*. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
  38. König J, Krahn U, Binder H. Visualizing the flow of evidence in network meta-analysis and characterizing mixed treatment comparisons. *Stat Med* 2013;32:5414–29.
  39. Higgins JP, Jackson D, Barrett JK, *et al*. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods* 2012;3:98–110.
  40. Peters JL, Sutton AJ, Jones DR, *et al*. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006;295:676.
  41. Chaimani A, Higgins JP, Mavridis D, *et al*. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013;8:e76654.
  42. Efthimiou O, Mavridis D, Cipriani A, *et al*. An approach for modelling multiple correlated outcomes in a network of interventions using odds ratios. *Stat Med* 2014;33:2275–87.
  43. Efthimiou O, Mavridis D, Riley RD, *et al*. Joint synthesis of multiple correlated outcomes in networks of interventions. *Biostatistics* 2015;16:84–97.
  44. Salanti G, Dias S, Welton NJ, *et al*. Evaluating novel agent effects in multiple-treatments meta-regression. *Stat Med* 2010;29:n/a–83.
  45. CINeMA: Confidence in Network Meta-Analysis [Software]. Institute of Social and Preventive Medicine, University of Bern. 2017.
  46. Hutton B, Wolfe D, Moher D, *et al*. Reporting guidance considerations from a statistical perspective: overview of tools to enhance the rigour of reporting of randomised trials and systematic reviews. *Evid Based Ment Health* 2017;20:46–52.
  47. Glue P, Donovan MR, Kolluri S, *et al*. Meta-analysis of relapse prevention antidepressant trials in depressive disorders. *Aust N Z J Psychiatry* 2010;44:697–705.
  48. Chaimani A, Salanti G, Leucht S, *et al*. Common pitfalls and mistakes in the set-up, analysis and interpretation of results in network meta-analysis: what clinicians should look for in a published article. *Evid Based Ment Health* 2017;20:88–94.