

# Initial treatment choices to achieve sustained response in major depression: a systematic review and network meta-analysis

Toshi A. Furukawa<sup>1</sup>, Kiyomi Shinohara<sup>1</sup>, Ethan Sahker<sup>1</sup>, Eirini Karyotaki<sup>2</sup>, Clara Miguel<sup>2</sup>, Marketa Ciharova<sup>2</sup>, Claudi L.H. Bockting<sup>3</sup>, Josefien J.F. Breedveld<sup>3</sup>, Aran Tajika<sup>1</sup>, Hissei Imai<sup>1</sup>, Edoardo G. Ostinelli<sup>4,5</sup>, Masatsugu Sakata<sup>1</sup>, Rie Toyomoto<sup>1</sup>, Sanae Kishimoto<sup>1</sup>, Masami Ito<sup>1</sup>, Yuki Furukawa<sup>6</sup>, Andrea Cipriani<sup>4,5</sup>, Steven D. Hollon<sup>7</sup>, Pim Cuijpers<sup>2</sup>

<sup>1</sup>Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health, Kyoto, Japan; <sup>2</sup>Department of Clinical, Neuro- and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit, Amsterdam, The Netherlands; <sup>3</sup>Department of Psychiatry & Centre for Urban Mental Health, University of Amsterdam, Amsterdam, The Netherlands; <sup>4</sup>Department of Psychiatry, Warneford Hospital, University of Oxford, Oxford, UK; <sup>5</sup>Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK; <sup>6</sup>Department of Neuropsychiatry, University of Tokyo Hospital, Tokyo, Japan; <sup>7</sup>Department of Psychology, Vanderbilt University, Nashville, TN, USA

*Major depression is often a relapsing disorder. It is therefore important to start its treatment with therapies that maximize the chance of not only getting the patients well but also keeping them well. We examined the associations between initial treatments and sustained response by conducting a network meta-analysis of randomized controlled trials (RCTs) in which adult patients with major depression were randomized to acute treatment with a psychotherapy (PSY), a protocolized antidepressant pharmacotherapy (PHA), their combination (COM), standard treatment in primary or secondary care (STD), or pill placebo, and were then followed up through a maintenance phase. By design, acute phase treatment could be continued into the maintenance phase, switched to another treatment or followed by discretionary treatment. We included 81 RCTs, with 13,722 participants. Sustained response was defined as responding to the acute treatment and subsequently having no depressive relapse through the maintenance phase (mean duration: 42.2±16.2 weeks, range 24-104 weeks). We extracted the data reported at the time point closest to 12 months. COM resulted in more sustained response than PHA, both when these treatments were continued into the maintenance phase (OR=2.52, 95% CI: 1.66-3.85) and when they were followed by discretionary treatment (OR=1.80, 95% CI: 1.21-2.67). The same applied to COM in comparison with STD (OR=2.90, 95% CI: 1.68-5.01 when COM was continued into the maintenance phase; OR=1.97, 95% CI: 1.51-2.58 when COM was followed by discretionary treatment). PSY also kept the patients well more often than PHA, both when these treatments were continued into the maintenance phase (OR=1.53, 95% CI: 1.00-2.35) and when they were followed by discretionary treatment (OR=1.66, 95% CI: 1.13-2.44). The same applied to PSY compared with STD (OR=1.76, 95% CI: 0.97-3.21 when PSY was continued into the maintenance phase; OR=1.83, 95% CI: 1.20-2.78 when PSY was followed by discretionary treatment). Given the average sustained response rate of 29% on STD, the advantages of PSY or COM over PHA or STD translated into risk differences ranging from 12 to 16 percentage points. We conclude that PSY and COM have more enduring effects than PHA. Clinical guidelines on the initial treatment choice for depression may need to be updated accordingly.*

**Key words:** Major depression, treatment choice, maintenance treatment, sustained response, psychotherapy, pharmacotherapy, combination therapy, cognitive behavioral therapy, network meta-analysis

**(World Psychiatry 2021;20:387–396)**

The two mainstays of acute treatment of major depression in adults are antidepressant medications and psychotherapies, each backed by several hundred randomized controlled trials<sup>1,2</sup>. After remission from the episode, it is also well documented that continuing pharmacotherapies<sup>3,4</sup> or psychotherapies<sup>5</sup>, or sequentially introducing psychotherapies as add-on to pharmacological treatments<sup>6</sup>, can reduce the depressive relapse rate in the maintenance phase.

Antidepressants are currently among the most frequently prescribed medications worldwide, being taken by 10% or more of the general population annually in some high-income countries<sup>7</sup>. More and more patients seem to be on longer-term antidepressant treatment: in the US, 44% of the current recipients had been on antidepressants for more than five years in 2015, compared with only 13% in 1996<sup>8</sup>.

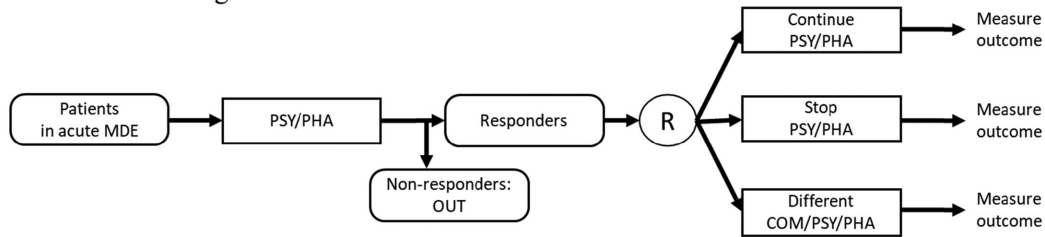
Three types of trial designs have been used in the literature to assess the efficacy of maintenance treatments in depression<sup>9</sup>. The most commonly used is the “enrichment design” (type A in Figure 1), in which patients who have responded to an acute treatment are subsequently randomized to various maintenance treatments. The second (type B) is the “continuation design”, in which patients with depression are randomly allocated to re-

ceive an intervention or a control and then the entire cohort is followed up into the maintenance phase. A variant of the latter is the “extension design” (type C), in which only participants who have responded to the acute treatment are followed up. In both type B and C studies, the follow-up maintenance therapy is by design the same as in the acute phase, or a new treatment, or is left to the therapist’s discretion in a naturalistic fashion.

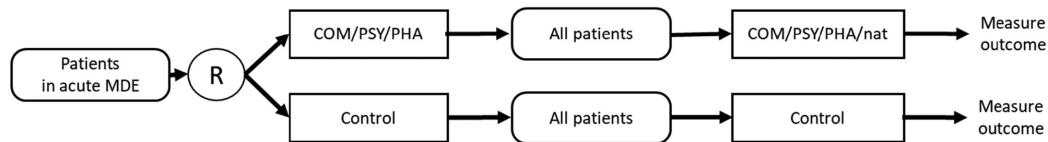
Systematic reviews of maintenance treatments to date have focused on type A trials to determine what should be done after successful acute treatment of depression<sup>3-6</sup>. While such information is clinically important, it cannot answer the clinically more pertinent question that faces every patient starting treatment for a depressive episode: “Which therapies can get me well and keep me well?”. Type A trials are enriched for, and therefore potentially biased in favor of, the first active therapy<sup>10,11</sup>. Only type B and C trials, in which randomization takes place at the beginning of the acute phase, can inform the initial treatment choice.

We hereby present the first systematic review and network meta-analysis (NMA) to determine which of the available therapies for depression chosen at the beginning of the acute phase are more likely to lead to sustained response in the maintenance phase. The NMA preserves the randomized structure of the evi-

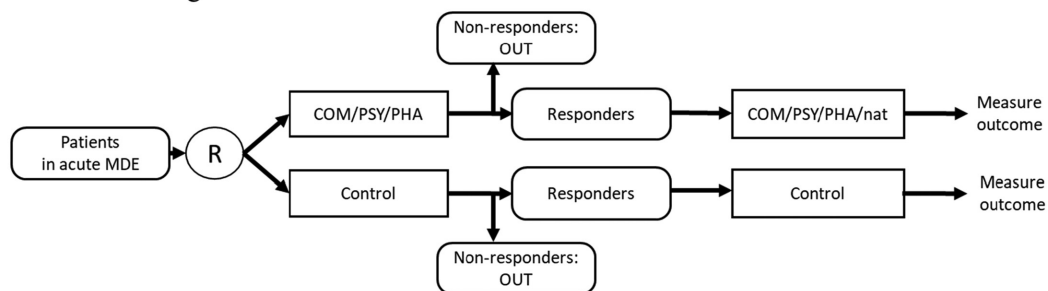
### A Enrichment design



### B Continuation design



### C Extension design



**Figure 1** Trial designs to examine maintenance treatment for depression. MDE – major depressive episode, COM – combination therapies, PHA – pharmacotherapies, PSY – psychotherapies, nat – discretionary treatment, R – randomization

dence network, i.e. treatment effects are first estimated separately for each study and then such study-specific estimates are synthesized for each treatment comparison and across the network, assuming constancy of the relative effect at each stage of the synthesis. This assumption of constancy is duly examined while conducting NMA.

## METHODS

We followed the PRISMA guideline for NMAs<sup>12</sup>. The protocol has been registered at the Open Science Framework (<https://osf.io/5qfuv/>).

### Data search

We identified relevant studies from three databases covering PubMed, EMBASE, PsycINFO, Cochrane Library, major trial registries, and regulatory agency websites. The first is a database of randomized trials of psychotherapies for depression, described at [www.osf.io/825c6](http://www.osf.io/825c6) and continuously updated<sup>13</sup> (the last search was conducted on January 1, 2020). The second is a database of

randomized trials of psychotherapies focusing on relapse prevention<sup>14</sup> (the last search was on October 13, 2019). The third is a database of randomized trials of antidepressant pharmacotherapies in relapse prevention<sup>9</sup> (the last search was on January 3-5, 2019). The search strings used in each database are provided in the supplementary information. Two independent raters judged the eligibility of the included studies.

### Study selection

We included randomized controlled trials in which any of the relevant interventions (see below) were compared with each other or with control conditions (see below) in the maintenance treatment of major depression, in type B or C studies (see Figure 1). We defined maintenance treatment as the continuation of treatment for six or more months. Because the distinction between a continuation phase to prevent relapses (re-emergence of the index episode) and a maintenance phase to prevent recurrences (appearance of a new episode)<sup>15</sup> is more theoretical than pragmatic<sup>3</sup>, we use the term maintenance therapy to refer to the longer-term treatment phase after the acute phase.

We included patients aged 18 years or older, of both genders,

with unipolar major depression diagnosed on the basis of standard operationalized criteria. We excluded studies that relied on a cutoff on a screening scale as an eligibility criterion and did not ascertain the diagnosis of depression. Studies in which 20% or more of the participants suffered from bipolar disorder, psychotic depression, treatment resistant depression or subthreshold depression were excluded. We also excluded RCTs which focused on patients with another concurrent primary psychiatric diagnosis or with a concomitant medical illness.

Among psychotherapies, we included any intervention involving “the informed and intentional application of clinical methods derived from established psychological principles to assist participants with their behaviors, cognitions and emotions, in directions that the participants deem desirable”<sup>16</sup>. Interventions could be delivered by any therapist, including psychiatrists, psychologists, nurses, social workers, and also lay health counsellors as long as they were trained to deliver the therapy, either in individual or group format, face-to-face or by Internet. We excluded unguided self-help interventions as they have been documented to be inferior to other delivery modalities for major depression<sup>17-19</sup>. Psychotherapies were further subcategorized into the following major types: cognitive behavioral therapy (CBT), behavioral activation therapy (BA), problem-solving therapy (PST), third-wave cognitive behavioral therapies (3W), interpersonal therapy (IPT), psychodynamic therapy (DYN), non-directive supportive therapy (SUP), and life review therapy (LRT)<sup>20-22</sup>.

Among pharmacotherapies, we included fixed or flexible dose regimens of antidepressants that have shown greater efficacy than placebo in acute treatment<sup>1</sup>. Only arms within the accepted dose ranges were included.

Controls included pill placebo; standard non-protocolized treatment in primary or secondary care, typically with pharmacotherapies (STD); and no treatment (NT) if the care as usual in the trial context involved virtually no intervention (operationally defined as less than one third of patients receiving any antidepressant).

The primary outcome was “sustained response”, defined as the proportion of patients who had responded in the acute treatment and who subsequently did not have depressive relapses during the maintenance phase. The proportion of sustained response, therefore, represented those who had responded to the acute phase treatment and maintained the response through the maintenance treatment, divided by the total number of patients randomized at the beginning of the acute phase treatment. We extracted the data reported at the time point closest to 12 months.

In some type B studies, when above-defined sustained response was not reported, we used the number of responders at the follow-up, either reported as dichotomous outcomes or imputed from the continuous outcomes using a validated imputation method<sup>23,24</sup>. We regarded all the dropouts as not showing sustained response. We examined the effect of this assumption by a sensitivity analysis limiting to studies with >90% follow-up.

The secondary outcome was all-cause discontinuation of

treatment, as a proxy measure of treatment acceptability. We had originally intended to also evaluate discontinuation due to adverse events (tolerability) and suicidality. However, too few studies reported these harm outcomes through the maintenance phase, and we present only narrative summaries for these outcomes.

## Data extraction and quality assessment

Two independent researchers extracted the data using a standardized form. Two independent raters assessed the risk of bias in included studies using Cochrane’s revised risk of bias tool for randomized trials<sup>25</sup>. We assessed the risk of bias for each comparison within the included studies referring to the primary outcome. Any disagreement between the two raters was resolved through discussion or in consultation with a third reviewer.

## Data synthesis and analysis

We evaluated psychotherapies (PSY), protocolized pharmacotherapies (PHA), and their combinations (COM), each of which could be continued into the maintenance treatment, switched to another treatment, or followed by discretionary treatment (nat). Controls were treatment as usual in primary or secondary care followed by the same discretionary treatment (STD), and pill placebo used through the acute and maintenance phase. Psychotherapies combined with protocolized pharmacotherapy or with non-protocolized primary or secondary care pharmacotherapy were counted towards COM. The influence of including the latter was examined in a sensitivity analysis.

We estimated the comparative efficacy and acceptability of these alternative treatments using the NMA methodology, by combining direct and indirect evidence for all relative treatment effects. We conducted contrast-based NMA to estimate odds ratios (ORs) with their 95% confidence intervals (CIs)<sup>26-28</sup>. Given the likely clinical and methodological heterogeneity among the included trials, we used the random effects model.

To examine the transitivity assumption that effect modifiers are distributed evenly across comparisons in the network (a primary requisite of NMA), we first made a table of important trial characteristics of the studies per comparison. We also examined transitivity statistically for the closed network by checking its consistency with the side-splitting test<sup>29</sup> and the design-by-treatment interaction test<sup>30</sup>. We evaluated the heterogeneity in the network with tau-squared in comparison with empirically derived evidence<sup>31</sup>. We further conducted a multivariate meta-regression analysis on age, proportion of women, baseline depression severity and total duration of treatment in order to examine if such factors affected constancy of ORs in the network.

We assessed small study effects, including publication bias, through visual inspection of the contour-enhanced funnel plot<sup>32</sup> and Egger’s test<sup>33</sup> of the aggregated pairwise comparisons between active interventions and control conditions.

We also performed several sensitivity analyses: a) limiting to studies which reported narrowly defined sustained response (see above); b) limiting to studies which followed up more than 90% of the randomized patients in all of their arms; c) limiting to studies in which the total duration of treatment was 12 months or longer; d) excluding studies at high risk of bias; e) excluding arms with non-protocolized primary or secondary care pharmacotherapy, because its contents may vary greatly; f) excluding arms with pill placebo, because they may change the nature of the trials<sup>34</sup>; and g) distinguishing all the subcategories of interventions or control conditions. We used CINeMA<sup>35</sup> to evaluate certainty of evidence for the network estimates.

The absolute benefits of the therapies were calculated from the ORs and the control event rate (CER) using the following formulae:  $RR=OR/(1-CER+OR*CER)$ ;  $EER=CER*RR$ ;  $RD=EER-CER$ , where RR is the relative risk, EER is the event rate in the intervention group, and RD is the risk difference (absolute benefit)<sup>36-38</sup>.

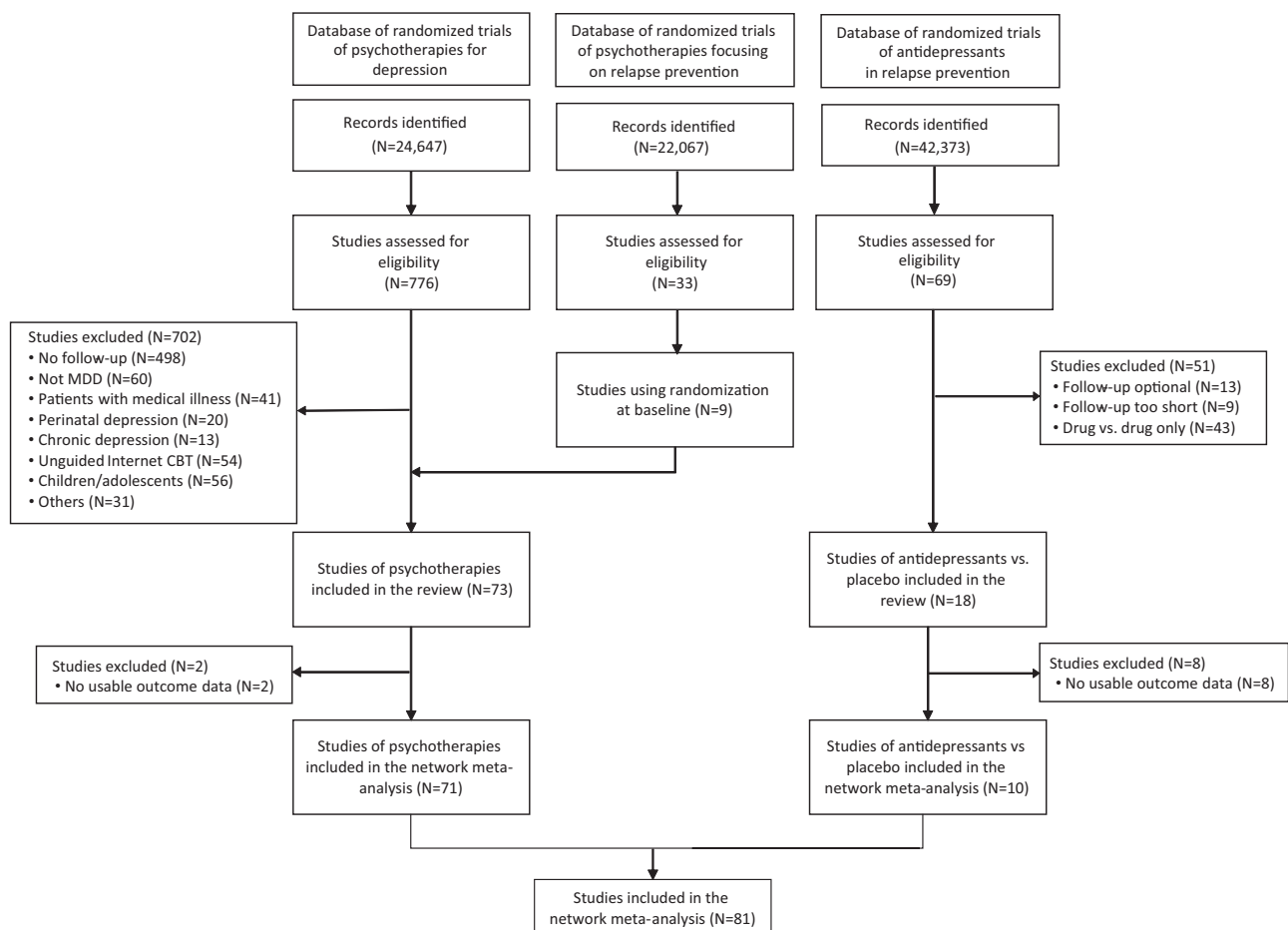
We employed the package netmeta 1.2-1 and dmetar 0.0.9 in R 4.0.3 (R Core Team, Vienna, Austria, 2020). Network meta-regressions were conducted with the network package<sup>39</sup> in STATA 16.1 (StataCorp, Texas, USA, 2020).

## RESULTS

### Studies selected and their characteristics

After examining 89,087 references in the three databases and 878 full text articles in detail, we included 81 studies (N=13,722). The PRISMA flow chart is presented in Figure 2. The references for the included trials and the reasons for exclusion of the others are provided in the supplementary information.

Table 1 summarizes the baseline characteristics of the included trials and their participants. The participants' weighted mean age (reported for 12,940 people) was  $43.4\pm 10.1$ , and 68% of the participants (8,668 out of 12,749 people for whom gender was reported) were women. The patients' baseline total score on the 17-item Hamilton Rating Scale for Depression<sup>40</sup> was  $21.8\pm 5.4$  in the 42 studies (N=7,918) that used this scale. The average total duration of treatment was  $42.2\pm 16.2$  weeks (range: 24-104 months) for the 81 studies. The average duration of the acute phase of treatment was  $10.4\pm 4.8$  weeks for 79 studies (two studies only provided the total length of acute plus maintenance phase and continued the same treatment through both phases). The weighted mean follow-up rate was 74.5%.



**Figure 2** PRISMA flow chart. MDD – major depressive disorder, CBT – cognitive behavioral therapy

**Table 1** Summary characteristics of the 81 included studies

Study design	
Type B	64
Type C	17
Number of arms (total=211)	
Two	44
Three	26
Four	10
Six	1
Publication year	
Earliest	1981
Median	2008
Latest	2019
Region	
North America	28
Europe	37
Asia	7
Cross-continental/Other	9
Randomization	
Individual	78
Cluster	3
Number of study centers	
Single	30
Multiple	51
Patient status	
Outpatients	59
Community	12
Inpatients	6
Others/Unclear	4
Treatment setting	
Community	11
Primary care	15
Secondary/Tertiary care	41
Others/Unclear	14
Diagnostic criteria	
DSM-5	2
DSM-IV	47
DSM-III-R	9
DSM-III	4
ICD-10	7
Research Diagnostic Criteria	8
Feighner criteria	4
Patients' gender, N women (%)	8,668/12,749 (68.0)
Patients' age (years, mean±SD)	43.4±10.1

**Table 1** (continued)

Depression baseline severity (mean±SD)	
HAMD-17 (42 studies)	21.8±5.4
BDI (8 studies)	24.9±7.6
BDI-II (7 studies)	26.8±9.3
Recurrent depression, % (32 studies)	62.6
Length of acute treatment (weeks, mean±SD) (79 studies)	10.4±4.8 (range: 4-30)
Length of total treatment (weeks, mean±SD) (81 studies)	42.2±16.2 (range: 24-104)
Follow-up rate, %	74.5

HAMD-17 – 17-item Hamilton Rating Scale for Depression, BDI – Beck Depression Inventory, BDI-II – Beck Depression Inventory, 2nd version

The 81 studies included 211 arms, which could be classified into 10 types and 34 subtypes of interventions. The most frequently examined intervention types included COM followed by naturalistic follow-up (COM→nat, 65 arms), PHA continued into the maintenance phase (PHA→PHA, 34 arms), PSY followed by naturalistic follow-up (PSY→nat, 30 arms), and treatment as usual in primary or secondary care through the acute and maintenance phase (STD, 25 arms).

The most frequently used types of psychotherapies in PSY and COM included CBT (59 arms), SUP (16 arms), IPT (11 arms), BA (8 arms), and DYN (7 arms). The most frequently used antidepressants were duloxetine (N=906 of 5,714 reported, 15.8%), agomelatine (N=836, 14.6%), paroxetine (N=644, 11.3%), venlafaxine (N=583, 10.2%) and fluoxetine (N=296, 5.2%).

Of the 155 comparisons, 40.6% were rated low for susceptibility bias, 49.4% for performance bias, 37.4% for attrition bias, 53.5% for assessment bias, and 1.3% for reporting bias. Overall, 89 (60.5%) were rated at high, 49 (33.3%) at moderate and 9 (6.1%) at low overall risk of bias.

## Network meta-analyses

Figure 3 presents the network of the interventions for the primary outcome. The nodes are well connected. Table 2 presents the network meta-analysis results for the primary outcome (sustained response) and the secondary outcome (all-cause discontinuation), and Figures 4 and 5 illustrate their ranked forest plots in comparison with STD.

COM brought about more sustained response than PHA, both if these treatments were continued into the maintenance phase (COM→COM vs. PHA→PHA: OR=2.52, 95% CI: 1.66-3.85) and if they were followed by discretionary treatment (COM→nat vs. PHA→nat: OR=1.80, 95% CI: 1.21-2.67). The same applied to COM when compared with standard therapy through the acute and maintenance phases (COM→COM vs. STD: OR=2.90, 95% CI: 1.68-5.01; COM→nat vs. STD: OR=1.97, 95% CI: 1.51-2.58)

(see Table 2 and Figure 4).

PSY was also more efficacious than PHA, both if these treatments were continued into the maintenance phase (PSY→PSY vs. PHA→PHA: OR=1.53, 95% CI: 1.00-2.35) and if they were followed by discretionary treatment (PSY→nat vs. PHA→nat: OR=1.66, 95% CI: 1.13-2.44). The same applied to PSY when compared with standard therapy through the acute and maintenance phases (PSY→PSY vs. STD: OR=1.76, 95% CI: 0.97-3.21; PSY→nat vs. STD: OR=1.83, 95% CI: 1.20-2.78) (see Table 2 and Figure 4).

PHA, continued or followed by discretionary treatment, did not differentiate from STD (PHA→PHA vs. STD: OR=1.15, 95% CI: 0.69-1.92; PHA→nat vs. STD: OR=1.10, 95% CI: 0.70-1.73) (see Table 2 and Figure 4).

Given the average sustained response rate on STD of 29% at 12 months (367 of 1,283 reported), the advantage (“absolute benefit”) of COM→nat over PHA→nat and STD would translate into a risk difference, respectively, of 14% (95% CI: 4 to 24%) and 16% (95% CI: 9 to 22%), while the advantage of PSY→nat over PHA→nat and STD can be calculated, respectively, as 12% (95% CI: 2 to 20%) and 14% (95% CI: 4 to 24%).

In terms of all-cause discontinuation, all the treatments appeared more acceptable than pill placebo. COM, PHA or PSY followed by discretionary treatment were generally as acceptable as STD. By contrast, stricter follow-up regimens, either by COM, PHA or PSY, tended to lead to more dropouts than STD (see Table 2 and Figure 5).

Transitivity of the network was preserved in terms of age, gender, and baseline depression severity. The global test of transitivity assumption was not suggestive of network inconsistency ( $p=0.98$ ); none of the side-splitting tests revealed inconsistency beyond chance. The common heterogeneity parameter tau-

squared was 0.196, within the empirically expected range for subjective outcomes for non-pharmacological interventions<sup>31</sup>. In network meta-regressions to examine sources of heterogeneity, age, proportion of women, baseline severity of depression and total duration of treatment, alone or in combination, did not show statistically significant effect modifications for any of the interventions. Funnel plots of active interventions against control conditions were not suggestive of small study effects ( $p=0.84$  and  $p=0.21$ , respectively).

The overall proportions of dropouts due to adverse events or suicidality through the long-term treatment were 10.3% (64 out of 619 reported in 6 studies) and 3.7% (29 out of 777 reported in 8 studies), respectively.

The sensitivity analyses sometimes had wide confidence intervals but generally produced results convergent with the primary analysis for sustained response. The results were more variable with regard to all-cause discontinuation (see supplementary information).

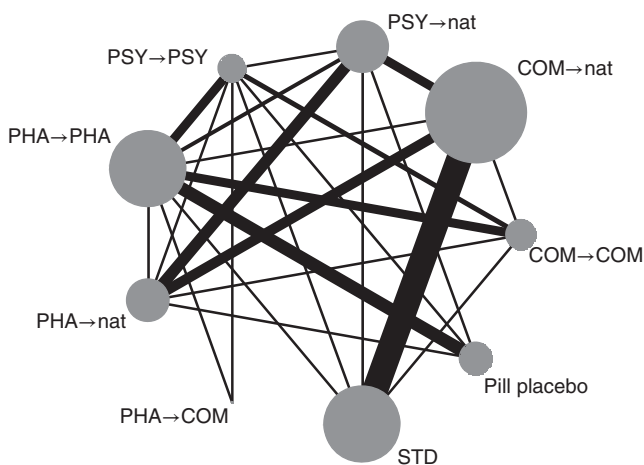
We also conducted NMA distinguishing all intervention subtypes. There was suggestive evidence that combining DYN, CBT, IPT or BA with antidepressant pharmacotherapy or treatment as usual led to more sustained response than STD. The same was true for CBT (either continued in the maintenance phase or followed by discretionary treatment), and for BA (followed by discretionary treatment) compared to STD (see Figure 6).

The certainty of evidence was rated as moderate for COM→COM and COM→nat vs. STD; as low for PSY→PSY and PSY→nat vs. STD; as low for PHA→PHA vs. STD, and as moderate for PHA→nat vs. STD. It was high only for COM→COM and COM→nat vs. pill placebo (see supplementary information).

## DISCUSSION

We conducted the first systematic review and network meta-analysis of the initial intervention choices for major depressive episodes aimed to maximize the chance of not only getting the patients well but also keeping them well. We identified 81 relevant studies (13,722 patients), which constituted a well-connected network of pharmacotherapies, psychotherapies and their combinations with little overall evidence of intransitivity, inconsistency, heterogeneity or publication bias. Various sensitivity analyses corroborated the primary findings.

There were two major findings of this study. First, acute phase combination therapies, either continued into the maintenance phase (COM→COM) or followed by discretionary treatment (COM→nat), outperformed both acute phase pharmacotherapies, continued or followed by discretionary treatment (PHA→PHA and PHA→nat), and standard therapy through the acute and maintenance phases (STD). Given the average sustained response rate of 29% on STD, the advantages of COM over PHA or STD translated into risk differences ranging from 14 to 16 percentage points. Second, psychotherapies, continued into the maintenance phase (PSY→PSY) or followed by discretionary treatment (PSY→nat), also outperformed pharmacotherapies and standard



**Figure 3** Network diagram for sustained response. COM - combination therapies, PHA - pharmacotherapies, PSY - psychotherapies, STD - standard treatment in primary or secondary care, nat - discretionary treatment. The size of the node is proportionate to the number of participants allocated to that node; the width of the line is proportionate to the number of studies examining that comparison.

**Table 2** Network meta-analyses for sustained response (efficacy) and all-cause discontinuation (acceptability) of various treatment modalities

COM→COM	<b>1.92</b> (1.04-3.54)	1.31 (0.68-2.51)	0.95 (0.57-1.58)	0.68 (0.43-1.07)	1.43 (0.75-2.75)	0.81 (0.19-3.55)	1.45 (0.78-2.68)	<b>0.33</b> (0.18-0.60)
1.47 (0.85-2.53)	COM→nat	0.68 (0.44-1.07)	<b>0.50</b> (0.25-0.97)	<b>0.35</b> (0.20-0.64)	0.75 (0.48-1.16)	0.42 (0.09-1.95)	0.76 (0.56-1.02)	<b>0.17</b> (0.08-0.35)
1.59 (0.91-2.76)	1.08 (0.74-1.56)	PSY→nat	0.73 (0.36-1.45)	<b>0.52</b> (0.28-0.94)	1.09 (0.70-1.70)	0.62 (0.13-2.88)	1.11 (0.68-1.81)	<b>0.25</b> (0.12-0.51)
<b>1.65</b> (1.04-2.61)	1.12 (0.62-2.03)	1.04 (0.57-1.88)	PSY→PSY	0.71 (0.45-1.14)	1.50 (0.74-3.04)	0.85 (0.20-3.57)	1.52 (0.78-2.99)	<b>0.34</b> (0.19-0.64)
<b>2.52</b> (1.66-3.85)	<b>1.72</b> (1.04-2.84)	1.59 (0.98-2.60)	<b>1.53</b> (1.00-2.35)	PHA→PHA	<b>2.11</b> (1.13-3.91)	1.20 (0.29-4.99)	<b>2.13</b> (1.19-3.84)	<b>0.48</b> (0.32-0.73)
<b>2.64</b> (1.46-4.76)	<b>1.80</b> (1.21-2.67)	<b>1.66</b> (1.13-2.44)	1.60 (0.85-3.02)	1.05 (0.61-1.81)	PHA→nat	0.57 (0.12-2.65)	1.01 (0.62-1.67)	<b>0.23</b> (0.11-0.48)
2.97 (0.71-12.45)	2.02 (0.46-8.79)	1.87 (0.43-8.13)	1.80 (0.45-7.26)	1.18 (0.29-4.76)	1.12 (0.25-4.98)	PHA→COM	1.78 (0.39-8.21)	0.40 (0.09-1.78)
<b>2.90</b> (1.68-5.01)	<b>1.97</b> (1.51-2.58)	<b>1.83</b> (1.20-2.78)	<b>1.76</b> (0.97-3.21)	1.15 (0.69-1.92)	1.10 (0.70-1.73)	0.98 (0.22-4.27)	STD	<b>0.23</b> (0.11-0.46)
<b>5.05</b> (3.00-8.51)	<b>3.44</b> (1.91-6.18)	<b>3.18</b> (1.79-5.66)	<b>3.06</b> (1.81-5.18)	<b>2.00</b> (1.47-2.73)	<b>1.91</b> (1.02-3.57)	1.70 (0.41-7.13)	1.74 (0.96-3.16)	Pill placebo

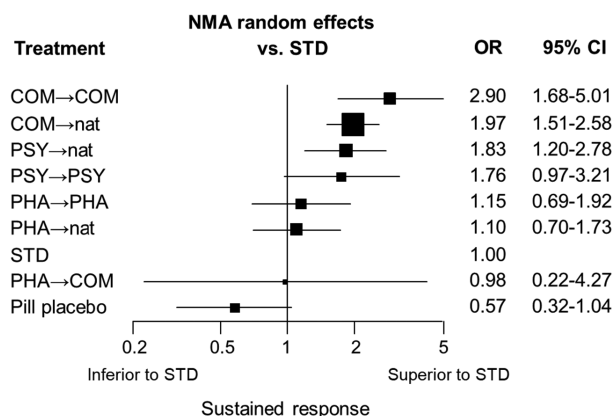
Values are odds ratios (ORs) with 95% confidence intervals. OR>1 in the lower-left half indicates that the treatment in the column is more effective than the treatment in the row. OR<1 in the upper-right half indicates that the treatment in the row is more acceptable than the treatment in the column. COM – combination therapies, PHA – pharmacotherapies, PSY – psychotherapies, STD – standard treatment in primary or secondary care, nat – discretionary treatment

therapy. The expected advantages were 12% for psychotherapies followed by discretionary treatment (PSY→nat) over the corresponding pharmacotherapies (PHA→nat), and 14% over STD.

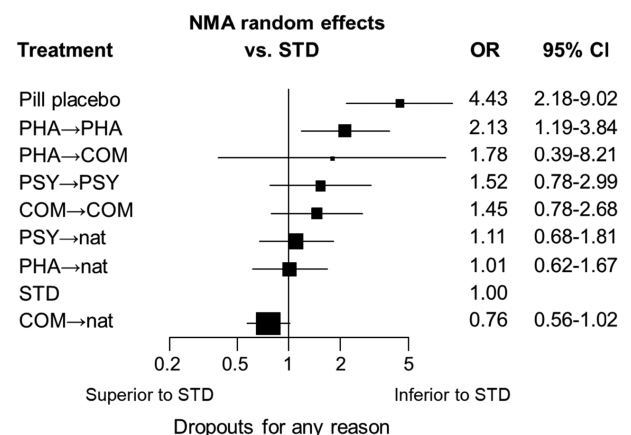
In the current systematic review, pharmacotherapies, while demonstrably superior to pill placebo, did not differentiate from standard treatment either if continued into the maintenance phase or followed by discretionary treatment.

This study provides strong answers to two long-held questions about psychotherapies<sup>11</sup>. First, it shows that the effects of acute phase psychotherapies are enduring. There was suspicion that, even when those responding to acute phase psychotherapies but

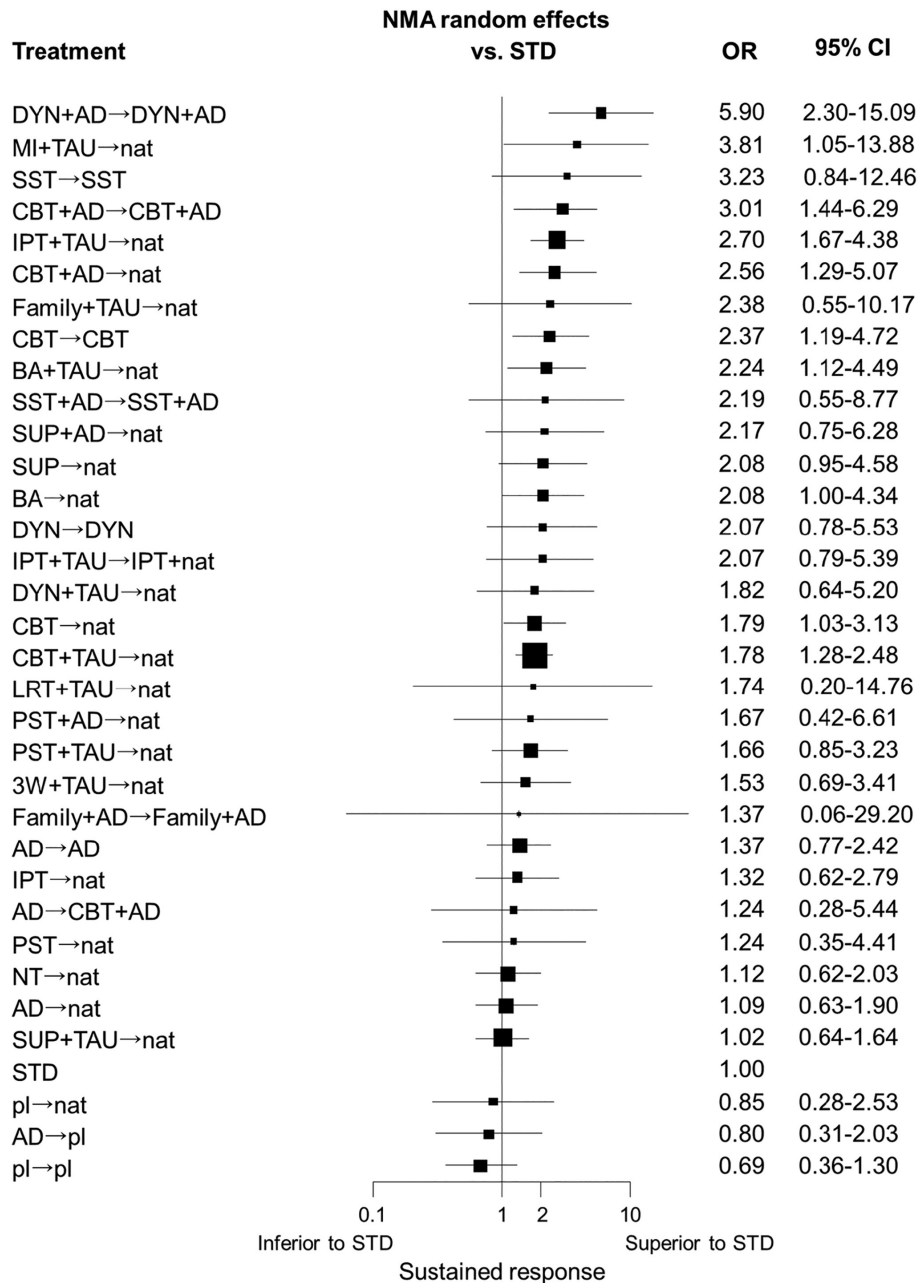
receiving no further psychotherapy did as well as those responding to acute phase pharmacotherapies and receiving maintenance pharmacotherapies<sup>5</sup>, this would not constitute proof that the acute effects of psychotherapies were enduring. The assumption was that those responding to acute phase psychotherapies may be systematically different from those responding to acute phase pharmacotherapies<sup>11,41</sup>. In this study, we only included trials that randomized participants into psychotherapies or pharmacotherapies at the beginning of the acute treatment and took these numbers as denominators in the analyses according to the intention-to-treat principle. The results clearly show that



**Figure 4** Ranked forest plot for sustained response. NMA – network meta-analysis, OR – odds ratio, CI – confidence interval, COM – combination therapies, PHA – pharmacotherapies, PSY – psychotherapies, STD – standard treatment in primary or secondary care, nat – discretionary treatment



**Figure 5** Ranked forest plot for all-cause discontinuation. NMA – network meta-analysis, OR – odds ratio, CI – confidence interval, COM – combination therapies, PHA – pharmacotherapies, PSY – psychotherapies, STD – standard treatment in primary or secondary care, nat – discretionary treatment



**Figure 6** Ranked forest plot for sustained response with intervention subtypes. NMA – network meta-analysis, OR – odds ratio, CI – confidence interval, STD – standard treatment in primary or secondary care, DYN – psychodynamic therapy, AD – protocolized antidepressant pharmacotherapy, MI – motivational interviewing, TAU – treatment as usual, nat – discretionary treatment, SST – social skills training, CBT – cognitive behavioral therapy, IPT – interpersonal therapy, Family – family therapy, BA – behavioral activation therapy, SUP – non-directive supportive therapy, LRT – life review therapy, PST – problem-solving therapy, 3W – third-wave cognitive behavioral therapy, NT – no treatment, pl – pill placebo

acute phase psychotherapies, even when not followed by maintenance psychotherapies, outperformed protocolized pharmacotherapies, standard treatment, and pill placebo.

Second, the findings suggest that adding pharmacotherapies does not interfere with the enduring effects of psychotherapies. The combination therapies followed by discretionary treatment were as effective as the corresponding psychotherapies (OR=1.08, 95% CI: 0.74-1.56), although the confidence intervals

are relatively wide and cannot completely exclude the interference hypothesis (according to which the OR should be smaller than 1.0)<sup>11,42</sup>.

The duration of total treatment ranged between 6 and 24 months. However, heterogeneity among the relative treatment effects was within empirically expected ranges<sup>31</sup>. Moreover, network meta-regression showed no evidence of an influence of the timing of the follow-up on ORs for any treatment comparisons.



A sensitivity analysis limiting studies to those in which the duration of treatment was 12 months or longer also produced similar results. It is therefore safe to assume that the obtained ORs for sustained response remain reasonably constant for total lengths of treatment ranging between 6 and 24 months. Such constancy of relative effect indices is in line with findings from pharmacological maintenance therapies for depression<sup>3</sup> and, more generally, across medical interventions<sup>36</sup>.

There are many types of psychotherapies and pharmacotherapies. While there is only limited evidence supporting differences within each category<sup>1,2</sup>, it would be helpful for clinical purposes to have insight as to which particular therapies are backed by stronger evidence. When we conducted the network meta-analysis for different subtypes of psychotherapies, there was consistent evidence that CBT (in combination or alone) and BA led to more sustained response than standard treatment. There were less consistent but similar trends for DYN and IPT. For other psychotherapies, there were too few studies and the corresponding confidence intervals were wide. With regard to pharmacotherapies, we were unable to examine the subtle differences among individual antidepressants in their ability to achieve sustained response. There were too many antidepressants used in the current network (hence too few patients for individual drugs) and several studies allowed use of several different antidepressants within their arms.

This study has several limitations. First, the maximum duration of the included trials was 24 months. The relative performance of the initial treatment choices if followed up for longer periods remains unknown. Second, many trials used a naturalistic follow-up after their protocolized acute treatment phase, and the exact content of treatment in the follow-up phase was seldom reported. Differences in this phase may have affected sustained response rates. However, such concerns are mitigated as the rankings among COM, PSY and PHA were similar when they were followed by discretionary treatment or when each was continued into the maintenance phase, as well as in a sensitivity analysis excluding trials using the discretionary follow-up.

Third, the weighted mean follow-up rate was 74.5%. The superiority of COM or PSY by 12-16% could be counterbalanced by whatever may have happened to the 25% who were lost to follow-up. However, a sensitivity analysis limiting to studies with 90% or greater follow-up confirmed the superiority of PSY and COM over STD. Fourth, only trials comparing PHA versus placebo could have been double-blind, which may have disadvantaged PHA in comparison with other treatments which were examined only in single-blind or open studies. The network without placebo-controlled trials, however, produced essentially similar efficacy estimates for all comparisons.

Fifth, the adverse effects of the available treatment choices were not well documented in the original studies and were therefore not amenable to systematic comparisons in the current network meta-analysis. Rare but critical events such as suicidality, and more common yet subtle downsides such as withdrawal symptoms from antidepressants should be more systematically measured and reported to appropriately inform our treatment choices<sup>43</sup>.

Lastly, we did not examine studies that randomized the remitted patients to completely new treatments after successful acute therapies<sup>6</sup>. Wisely sequencing different treatments has a potential to perform even better than simply choosing the best initial treatment<sup>44-46</sup>.

## CONCLUSIONS

Initiating the treatment of a major depressive episode with combination therapies or psychotherapies alone may lead to 12-16% increments in rates of sustained response at one year, relative to protocolized pharmacotherapies or standard treatment in primary or secondary care. Psychotherapies with the greatest support for such superiority include CBT, BA, and to a lesser degree DYN and IPT. Patients and their therapists may be well advised to seriously consider these psychotherapies as their initial treatment choices. However, availability and affordability of quality psychotherapies may be a major obstacle<sup>47-49</sup>.

Combining psychotherapies with pharmacotherapies has an edge in terms of sustained response but has risks of side effects and potential withdrawal symptoms. Such combinations may be reserved for those who value faster relief or who may be deemed difficult to treat<sup>22</sup>. Others may wish to consider them as sequenced treatments when initial therapies fail.

Findings from this study are robust enough to put the currently dominant practices relying on antidepressants into perspective, especially in the context of increasingly prevalent and protracted prescriptions<sup>7,8</sup>. Clinical guidelines may need to be updated accordingly. We also call for appropriately designed and adequately powered studies that examine alternative and sequential strategies to both get patients well and keep them well. Such studies need to consider cost-effectiveness and monitor suicidality and withdrawal symptoms systematically.

## ACKNOWLEDGEMENTS

This study was supported by the Japan Society for the Promotion of Science (grant no. 17K19808). E.G. Ostinelli is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, and by the NIHR Oxford Health Biomedical Research Centre (grant no. BRC-1215-20005). A. Cipriani is supported by the NIHR Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant no. RP-2017-08-ST2-006), by the NIHR Oxford and Thames Valley Applied Research Collaboration, and by the NIHR Oxford Health Biomedical Research Centre (grant no. BRC-1215-20005). Supplementary information on the study is available at <https://osf.io/5qfuw/>.

## REFERENCES

1. Cipriani A, Furukawa 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.
2. Cuijpers P, Quero S, Noma H et al. Psychotherapies for depression: a network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry* 2021;20:283-93.
3. 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.

4. 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* 2015;19:pyv076.
5. Cuijpers P, Hollon SD, van Straten A et al. Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis. *BMJ Open* 2013;3:e002542.
6. Breedvelt JJE, Brouwer ME, Harrer M et al. Psychological interventions as an alternative and add-on to antidepressant medication to prevent depressive relapse: systematic review and meta-analysis. *Br J Psychiatry* (in press).
7. Jorm AF, Patten SB, Brugha TS et al. Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries. *World Psychiatry* 2017;16:90-9.
8. Luo Y, Kataoka Y, Ostinelli EG et al. National prescription patterns of antidepressants in the treatment of adults with major depression in the US between 1996 and 2015: a population representative survey based analysis. *Front Psychiatry* 2020;11:35.
9. 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.
10. Furukawa TA, Miura T, Chaimani A et al. Using the contribution matrix to evaluate complex study limitations in a network meta-analysis: a case study of bipolar maintenance pharmacotherapy review. *BMC Res Notes* 2016;9:218.
11. Hollon SD. Is cognitive therapy enduring or antidepressant medications iatrogenic? Depression as an evolved adaptation. *Am Psychol* 2020;75:1207-18.
12. Hutton B, Salanti G, Caldwell DM et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162:777-84.
13. Cuijpers P, Karyotaki E, Ciharova M. A meta-analytic database of randomised trials on psychotherapies for depression. [www.osf.io/825c6](http://www.osf.io/825c6).
14. Breedvelt JJE, Warren FC, Brouwer ME et al. Individual participant data (IPD) meta-analysis of psychological relapse prevention interventions versus control for patients in remission from depression: a protocol. *BMJ Open* 2020;10:e034158.
15. Kupfer DJ. Long-term treatment of depression. *J Clin Psychiatry* 1991;52 (Suppl.):28-34.
16. Campbell LF, Norcross JC, Vasquez MJ et al. Recognition of psychotherapy effectiveness: the APA resolution. *Psychotherapy* 2013;50:98-101.
17. Cuijpers P, Noma H, Karyotaki E et al. Effectiveness and acceptability of cognitive behavior therapy delivery formats in adults with depression: a network meta-analysis. *JAMA Psychiatry* 2019;76:700-7.
18. Karyotaki E, Efthimiou O, Miguel C et al. Internet-based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. *JAMA Psychiatry* 2021;78:361-71.
19. Furukawa TA, Sukanuma A, Ostinelli EG et al. Dismantling, optimising and personalising internet cognitive-behavioural therapy for depression: a systematic review and component network meta-analysis using individual participant data. *Lancet Psychiatry* 2021;8:500-11.
20. Cuijpers P, van Straten A, Andersson G et al. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *J Consult Clin Psychol* 2008;76:909-22.
21. Cuijpers P, Karyotaki E, de Wit L et al. The effects of fifteen evidence-supported therapies for adult depression: a meta-analytic review. *Psychother Res* 2020;30:279-93.
22. Cuijpers P, Noma H, Karyotaki E et al. A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. *World Psychiatry* 2020;19:92-107.
23. Furukawa TA, Cipriani A, Barbui C et al. Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol* 2005;20:49-52.
24. da Costa BR, Rutjes AW, Johnston BC et al. Methods to convert continuous outcomes into odds ratios of treatment response and numbers needed to treat: meta-epidemiological study. *Int J Epidemiol* 2012;41:1445-59.
25. Sterne JAC, Savovic J, Page MJ et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
26. Bakbergenuly I, Hoaglin DC, Kulinskaya E. Pitfalls of using the risk ratio in meta-analysis. *Research Synthesis Methods* 2019;10:398-419.
27. Doi SA, Furuya-Kanamori L, Xu C et al. Questionable utility of the relative risk in clinical research: a call for change to practice. *J Clin Epidemiol* (in press).
28. White IR, Turner RM, Karahalios A et al. A comparison of arm-based and contrast-based models for network meta-analysis. *Stat Med* 2019;38:5197-213.
29. Dias S, Welton NJ, Caldwell DM et al. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932-44.
30. 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.
31. Turner RM, Davey J, Clarke MJ et al. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *Int J Epidemiol* 2012;41:818-27.
32. Peters JL, Sutton AJ, Jones DR et al. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol* 2008;61:991-6.
33. Egger M, Davey Smith G, Schneider M et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-34.
34. Salanti G, Chaimani A, Furukawa TA et al. Impact of placebo arms on outcomes in antidepressant trials: systematic review and meta-regression analysis. *Int J Epidemiol* 2018;47:1454-64.
35. Papakonstantinou T, Nikolakopoulou A, Higgins JPT et al. CINeMA: software for semiautomated assessment of the confidence in the results of network meta-analysis. *Campbell Syst Rev* 2020;16:e1080.
36. Furukawa TA, Guyatt GH, Griffith LE. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. *Int J Epidemiol* 2002;31:72-6.
37. Alhazzani W, Walter SD, Jaeschke R et al. Does treatment lower risk? Understanding the results. In: Guyatt G, Rennie D, Meade MO et al (eds). *Users' guides to the medical literature: a manual for evidence-based clinical practice*, 3rd ed. New York: McGraw-Hill, 2014:87-93.
38. Higgins JP, Thomas J (eds). *Cochrane handbook for systematic reviews of interventions*, version 6. <https://training.cochrane.org/handbook/current>.
39. White IR. Network meta-analysis. *Stata J* 2015;15:951-85.
40. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56-62.
41. Klein DF. Preventing hung juries about therapy studies. *J Consult Clin Psychol* 1996;64:81-7.
42. DeRubeis RJ, Zajecka J, Shelton RC et al. Prevention of recurrence after recovery from a major depressive episode with antidepressant medication alone or in combination with cognitive behavioral therapy: phase 2 of a 2-phase randomized clinical trial. *JAMA Psychiatry* 2020;77:237-45.
43. Cuijpers P. Targets and outcomes of psychotherapies for mental disorders: an overview. *World Psychiatry* 2019;18:276-85.
44. Guidi J, Fava GA. Sequential combination of pharmacotherapy and psychotherapy in major depressive disorder: a systematic review and meta-analysis. *JAMA Psychiatry* 2021;78:261-9.
45. Lavori PW, Dawson R. Dynamic treatment regimes: practical design considerations. *Clin Trials* 2004;1:9-20.
46. Breedvelt JJE, Warren FC, Segal Z et al. Continuation of antidepressants vs sequential psychological interventions to prevent relapse in depression. An individual participant data meta-analysis. *JAMA Psychiatry* (in press).
47. Hepner KA, Greenwood GL, Azocar F et al. Usual care psychotherapy for depression in a large managed behavioral health organization. *Adm Policy Ment Health* 2010;37:270-8.
48. van Ommeren M. Targets and outcomes of psychological interventions: implications for guidelines and policy. *World Psychiatry* 2019;18:295-6.
49. Jarrett RB. Can we help more? *World Psychiatry* 2020;19:246-7.

DOI:10.1002/wps.20906