



Comparative efficacy and tolerability of pharmacological and somatic interventions in adult patients with treatment-resistant depression: a systematic review and network meta-analysis

Katerina Papadimitropoulou, Carla Vossen, Andreas Karabis, Christina Donatti & Nicole Kubitz

To cite this article: Katerina Papadimitropoulou, Carla Vossen, Andreas Karabis, Christina Donatti & Nicole Kubitz (2017) Comparative efficacy and tolerability of pharmacological and somatic interventions in adult patients with treatment-resistant depression: a systematic review and network meta-analysis, Current Medical Research and Opinion, 33:4, 701-711, DOI: 10.1080/03007995.2016.1277201

To link to this article: <https://doi.org/10.1080/03007995.2016.1277201>

 View supplementary material [↗](#)

 Accepted author version posted online: 30 Dec 2016.
Published online: 06 Feb 2017.

 Submit your article to this journal [↗](#)

 Article views: 677

 View related articles [↗](#)

 View Crossmark data [↗](#)

 Citing articles: 4 View citing articles [↗](#)

ORIGINAL ARTICLE

Comparative efficacy and tolerability of pharmacological and somatic interventions in adult patients with treatment-resistant depression: a systematic review and network meta-analysis

Katerina Papadimitropoulou^a, Carla Vossen^a, Andreas Karabis^a, Christina Donatti^b and Nicole Kubitz^c

^aMapi Group, Houten, Netherlands; ^bJanssen-Cilag UK, High Wycombe, United Kingdom; ^cJanssen-Cilag GmbH, Neuss, Germany

ABSTRACT

Objective: Major depressive disorder (MDD) affects about 10–15% of the general population in a lifetime. A considerable number of patients fail to achieve full symptom remission despite adequate treatment and are considered treatment resistant (TRD). The current study compared the relative efficacy and tolerability of pharmacological and somatic TRD interventions by means of a Bayesian network meta-analysis.

Research design and methods: An electronic literature search of MEDLINE, MEDLINE In-Process, EMBASE, PsycInfo, EconLit and Cochrane Library databases for trials published between September 2003 and September 2014 was conducted. Key outcomes extracted were disease severity change from baseline, response and remission rates at various timepoints and discontinuation due to adverse events.

Results: Of the 3876 abstracts identified, 31 publications/randomised controlled trials (RCTs) were included in the analysis; 19 RCTs investigating 13 pharmacological interventions and 12 RCTs investigating electroconvulsive therapy (ECT) and repetitive transcranial magnetic stimulation (rTMS). The evidence synthesis investigating efficacy outcomes of TRD treatments demonstrated superior efficacy for ketamine compared to pharmacological and somatic interventions at 2 weeks after treatment initiation. At 4, 6 and 8 weeks, quetiapine augmentation (800 mg/day) and risperidone augmentation were found to be the first and second best treatments, respectively. Networks were small for response rate and remission rate outcomes at most timepoints. The most tolerable treatment was lamotrigine augmentation showing a comparable profile to placebo/sham.

Conclusions: This analysis revealed scarcity of long-term data on sustained remission that would allow a comparative long-term efficacy assessment. Key limitations of the analysis can be considered the search timeframe and the use of mapping formula for the depression scores.

ARTICLE HISTORY

Received 7 September 2016
Revised 23 December 2016
Accepted 23 December 2016

KEYWORDS

Antidepressant effect; Bayesian hierarchical model; Disease severity; Network meta-analysis; Ranking; Treatment-resistant depression

Introduction

Major depressive disorder (MDD), also known as unipolar depression, is a mental disorder mainly characterized by depressed mood and/or loss of interest or pleasure in daily activities for more than 2 weeks¹. In a lifetime, MDD affects about 15% of the general population in high-income countries and about 11% of the general population in low- to middle-income countries². MDD is associated with significant morbidity and mortality. The Global Burden of Disease study³ has quantified the burden of non-fatal health outcomes for 289 diseases and injuries between 1990–2010. This is done by determining the years of life lived in less than ideal health (years lived with disability [YLDs]). Of the 777,491,000 total global YLDs in 2010, 63 million (8%) were estimated to be caused by MDD, thereby ranking second in the top 25 causes of global YLDs³.

Treatment options for MDD include pharmacological and somatic/non-pharmacological interventions. Selective serotonin reuptake inhibitors (SSRIs) are the current first choice

of antidepressant drugs, followed by serotonin and noradrenaline reuptake inhibitors (SNRIs). Other classes of antidepressants include monoamine oxidase inhibitors (MAOIs) and tricyclic antidepressants (TCAs)⁴. Somatic/non-pharmacological strategies include psychotherapy (often in conjunction with pharmacotherapy), electroconvulsive therapy (ECT; often used as late-line treatment) and transcranial magnetic stimulation (TMS), among others⁵. The desired goal of treatment for depression is full symptom remission. However, not all MDD patients achieve this outcome; the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial reported a theoretical cumulative remission rate after four acute treatment steps of 67%⁶. Patients who require more treatment steps are considered treatment/therapy resistant, a concept initially proposed by Lehmann⁷ and Sartorius N.⁸ in their seminal studies of classifying depressive disorders. In clinical practice, there is no consensus on an operational definition of treatment-resistant depression (TRD) and although several staging models exist, they are not routinely used.

CONTACT Katerina Papadimitropoulou ✉ kpapadimitropoulou@mapiigroup.com Mapi Group, De Molen 84, 3995 AX, Houten, The Netherlands

 Supplemental data for this article can be accessed [here](#).

The Thase and Rush Staging Model (TSRM)⁹ suggests that patients are staged according to the number of failures of antidepressant classes with treatment resistance increasing when failure is reported after TCA, MAOI and ECT trials. In general, failure to respond to two or more different antidepressants prescribed at adequate dose and duration is one of the most commonly proposed definitions. It is also accepted in a regulatory setting^{1,9-11} and was the definition adopted for this study.

TRD pharmacological treatment can follow different strategies: optimization of the antidepressant dose, “switching” within and between classes of compounds, or “combining” antidepressant therapies with each other or with, for example, antipsychotic therapies (augmentation/adjunctive therapies). Most of the current interventions are not specifically approved for use in TRD. In the United States and Mexico, one product is approved for the treatment of TRD: olanzapine in a fixed dose combination with fluoxetine, marketed as Symbyax¹². In clinical practice, ECT is used to treat patients with TRD or patients with depression who are judged to be at significant risk for suicide, harm to others or severe self-neglect whereas availability and use of repetitive transcranial magnetic stimulation (rTMS) seems more limited and restricted¹³.

Several studies utilizing pairwise meta-analyses have investigated the treatment effect of a class of treatments in MDD and TRD – atypical antipsychotics¹⁴⁻¹⁶, SSRIs¹⁷, lithium augmentation¹⁸ and rTMS¹⁹. Recently, a network meta-analysis comparing treatment effects of all augmentation agents in TRD was published²⁰. The effect of different doses of adjunctive atypical antipsychotics was explored in a subsequent network meta-analysis (NMA)²¹.

The objective of this study was to compare the efficacy and the tolerability of TRD treatments in adult patients by means of a Bayesian network meta-analysis. This study is the first to indirectly compare and rank pharmacological and somatic interventions.

Methods

Systematic literature review

This systematic literature review (SLR) was performed as part of a larger search aiming to identify, more broadly, the clinical and economic burden of TRD. The search for the efficacy and tolerability evidence of TRD treatments was thus combined with the search for clinical and economic burden and the identified literature was screened simultaneously. Only results from the evidence on the efficacy and tolerability of TRD treatments are described here; results on clinical and economic burden of the disease will be reported elsewhere. MEDLINE, MEDLINE In-Process, EMBASE, PsycInfo, EconLit (through OVID) and Cochrane Library databases (including CENTRAL, CDSR, CMR, DARE, HTAD, and NHS EED) were searched for relevant trials in the combined search. Predefined search strategies were used, using terms relevant to TRD and major depression, randomised controlled trials (RCTs) and the interventions/comparators of interest (available in Supplementary eTable 1 and eTable2). In addition,

clinicaltrials.gov was searched to confirm that all clinical trials of interest were captured. Additional studies were searched in the reference lists of all identified meta-analyses and systematic reviews. All searches were restricted to English language. OVID was initially searched on 21 October 2013 for literature published between 2003 and the date of search. A targeted update, that is, only for evidence on the efficacy and safety of TRD treatments, was performed September 2014 using additional search terms.

The relevance of each identified citation was assessed based on the title and abstract according to predefined selection criteria (details in Supplementary eTable 3). For the abstracts that met the selection criteria, the full-text publications were retrieved and evaluated for inclusion. The abstract and full-text selection process was performed by two researchers: the first researcher was screening and the second was checking the decision. Any disagreements were resolved by a third reviewer. Only RCTs including at least 10 patients per treatment arm were considered of interest as smaller RCTs are mostly exploratory in nature and lack precision. The patient population was defined as adult MDD patients who failed to respond to ≥ 2 antidepressant treatment regimens prescribed at adequate dose and duration. Interventions (and comparators) of interest were: SSRIs, SNRIs, TCAs, tetracyclic antidepressants (TeCAs), MAOIs, atypical antidepressants, antipsychotics, olanzapine/fluoxetine combination (OFC), adjunctive use of lithium, triiodothyronine (T3), lamotrigine, ketamine, ECT and repetitive transcranial magnetic stimulation (rTMS).

Outcome measures

Outcomes of interest were: disease severity change from baseline (CFB) measured on the Hamilton (HAM-D) or Montgomery-Åsberg depression rating scales (MADRS) or other depression rating scales, response, remission, relapse and recurrence rates, time to response or relapse and tolerability outcomes. Previous studies have shown that relatively small differences in improvement scores can produce relatively large differences in expected response rates²². Therefore, mean changes from baseline in MADRS scale scores at 2, 4, 6 and 8 weeks after baseline were selected as the primary outcome. Note that not all assessment time-points were available for all treatments. The majority of studies reported mean CFB on the MADRS scale, yet when studies reported depressive scores based on the 17-item HAM-D scale, a mapping formula to convert HAM-D₁₇ to MADRS scores (and their respective standard deviations)²³ was used: $MADRS = 1.04 \times HAM - D_{17} + 10.13$.

Secondary efficacy outcomes were response rates and remission rates. Response was defined as a reduction in MADRS scores from baseline of at least 50% at study endpoint. Remission was mostly defined as a MADRS score of 7 or lower at study endpoint. For studies that did not report the number or proportion of responders/remitters, an imputation method proposed and validated by Fukurawa *et al.*²⁴ was implemented. The same method has been previously used in other multi-treatment meta-analyses in MDD²⁵.

Tolerability was defined as the proportion of patients who withdrew during the study due to adverse events.

Data extraction and quality assessment

The relevance of each citation identified was assessed in a two-tiered approach: first titles and abstracts were screened for eligibility (by one researcher and checked by a second), and for those fulfilling the selection criteria, full texts were retrieved and assessed. If there was a disagreement between the two reviewers, a third reviewer was consulted. Full text articles and abstracts that met the inclusion criteria were included for data extraction. In addition, each study was critically appraised using the Cochrane Collaboration's tool for assessing risk of bias²⁶.

Network meta-analysis

The relative efficacy and tolerability of TRD treatments for the selected outcomes were evaluated using a Bayesian hierarchical NMA²⁷. In this analysis, a linear model with normal likelihood distribution was used for continuous outcomes, and a binomial likelihood with a logit link for dichotomous outcomes. For continuous outcomes, i.e. change in MADRS total score from baseline, pooled estimates of mean differences in change from baseline (Δ CFB) between two treatments with 95% credible intervals (CrIs) were calculated, with negative values indicating greater symptomatic relief. When studies reported CFB without the associated sampling variance, standard errors were calculated based on the available data. For categorical outcomes, i.e. response rates, remission rates and withdrawals due to adverse events, odds ratios (OR) with 95% CrIs were calculated. ORs above 1 for efficacy outcomes indicate better efficacy for a treatment relative to the comparator. In contrast, ORs above 1 for withdrawals due to adverse events indicate a worse outcome for a treatment relative to a comparator.

Non-informative prior distributions were assumed for all outcomes. In the presence of non-informative priors, CrIs can be interpreted similarly to confidence intervals using a frequentist approach. Moreover, if the 95% CrIs do not include 0 (for continuous outcomes) and 1 (for categorical outcomes), respectively, results can be considered statistically significant²⁸. Prior distributions of the relative treatment effects were assumed to be normal, with zero as a mean and a variance of 10,000, while a uniform distribution with range zero to five was used as the prior of the between-study standard deviation.

It can be expected that there is always some variation in patient characteristics, study sites and settings across studies; if these characteristics are effect modifiers of the relative treatment effects of interest, there will be heterogeneity in the evidence base^{29–32}. To allow for heterogeneity between studies, random effects models were evaluated. Random effects models assume that treatment effects may vary between studies, but come from a common distribution of treatment effects, with a mean for each treatment effect and a common between-study covariance matrix. Furthermore, to

address potential bias in our study, a number of scenario analyses were defined a priori. For each outcome, fixed and random effects models were evaluated, and the better fitting model was selected based on the deviance information criterion (DIC) which adds a penalty term, equal to the number of effective parameters³³. The posterior densities for unknown parameters were estimated using Markov chain Monte Carlo (MCMC) simulations. The results presented herein were based on 80,000 iterations on two chains, with a burn-in of 20,000 iterations. Convergence was assessed by visual inspection of trace plots. The accuracy of the posterior estimates was assessed using the Monte Carlo error for each parameter (Monte Carlo error <1% of the posterior standard deviation). The unrelated mean effects (UME) model was used to assess the consistency assumption, in which each of the treatment contrasts (for which evidence is available) represents a separate, unrelated basic parameter to be estimated. All models were implemented using the OpenBUGS version 3.2.2 (MRC Biostatistics Unit, Cambridge, UK) and RStudio (R version 3.1.2) and were based on the models defined by Dias *et al.*^{34,35}.

The Bayesian NMA provided posterior distributions of the relative treatment effects between interventions and the probability that one treatment is better than another for each outcome of interest. This probability is calculated based on the proportion of Markov chain Monte Carlo cycles in which the specific treatment estimate is better than the comparator³⁵. Ranking probabilities were reported as surface under the cumulative ranking curve (SUCRA), the inversely scaled average rank of treatment. SUCRA values range from 0 to 1; SUCRA is 1 when a treatment is certain to be the best and 0 when a treatment is certain to be the worst.

Results

Evidence base

A total of 3876 abstracts were retrieved and screened (Supplementary eFigure 1). After screening, 55 publications were eligible for data extraction.

Risk of bias assessment

The results of the critical appraisal of the identified studies (per study and in total) using the Cochrane risk of bias tool are shown in Supplementary eTable 4 and Supplementary eFigure 2 for the pharmacological studies and in Supplementary eTable 5 and Supplementary eFigure 3 for the somatic intervention studies.

Network meta-analysis feasibility assessment

Based on the extracted data, the feasibility of an NMA was assessed following the steps recommended by Cope *et al.*³⁶. First, a connected network of studies was formed, including 31 studies (Supplementary eFigure 4), of which 19 investigated 13 pharmacological interventions^{10,37–53} and 12 investigated ECT and rTMS^{54–65}.

As a second step, study and patient characteristics across studies that were a priori identified by clinical experts as potential effect modifiers were investigated (see Supplementary material for an overview of study characteristics of pharmacological and somatic intervention studies). The list of patient characteristics identified as potential treatment effect modifiers included: age at baseline; concomitant treatment for depression; mean symptom severity at baseline; duration of the current episode; substance abuse; presence of anxiety; socio-economic status; and prior treatment for depression (number and type).

Median age at baseline was similar across studies (Supplementary eFigure 5 and eFigure 6); in pharmacological intervention studies age ranged from 41 to 52 years (excluding an outlier study⁴⁸) and in the somatic intervention studies from 38 to 58 years (excluding an outlier study⁶⁶).

Eleven studies provided information on concomitant medication (Supplementary eTable 6). The most commonly used concomitant medications across studies were SSRIs and SNRIs. Excluding studies allowing concomitant medication resulted in a disconnected network, so potential modification of the treatment could not be further investigated.

Thirteen pharmacological and 10 somatic intervention studies reported baseline disease severity scores on various depression scales. To facilitate valid comparisons, MADRS scores were compared (using original or converted scores). The mean baseline MADRS score in pharmacological studies was 29.8 (range 25.2 to 33.7) compared to 34.1 for somatic intervention studies (range 18.9 to 43.5). The effect of this difference was explored in a scenario analysis separating the somatic intervention from the pharmacological ones. The mean duration of the current depressive episode was 25 months (range 5.8 to 48.5) in 10 pharmacological intervention studies and 20 months (range 4.5 to 39.9) in 8 somatic intervention studies. This patient characteristic was deemed comparable across studies therefore no additional analyses were performed.

Substance abuse was an exclusion criterion for most studies or was not reported (see Supplementary eTable 7 and eTable 8); hence, no additional analyses were performed. Only few studies reported the presence of anxiety, socioeconomic status or prior treatment for depression (Supplementary material). Therefore, no additional analyses could be performed.

In general, study design, inclusion criteria and baseline patient characteristics were deemed comparable across the studies of the network, leading to the conclusion that performing a valid NMA was feasible.

A key objective of this network meta-analysis was to compare pharmacological and somatic interventions via a common comparator, i.e. a placebo/sham arm by pooling placebo arms from pharmacological intervention studies with sham arms of somatic intervention studies. Our assessment of the validity of pooling these placebo/sham arms was based on the comparison of baseline disease severity scores (measured by the MADRS), the change in total score from baseline (measured by the MADRS), response rate, defined as 50% reduction on MADRS or HAM-D score from baseline and remission rate (defined as MADRS ≤ 10 or HAM-D score ≤ 7)

between studies. In total, 9 pharmacological intervention studies were placebo-controlled trials and 11 somatic intervention studies were sham-controlled and included in the assessment. On average, patients randomized in the placebo arms of the pharmacological intervention studies had lower MADRS scores at baseline (range 25.9 to 36.2) than patients randomized to sham arms (range 18.9 to 40.3). The average MADRS CFB at 2 weeks was -4.5 for placebo arms and -4.9 for sham arms with a similar pattern at other timepoints. Response rates were more variable across studies: placebo arm response rates were reported in seven studies and ranged between 18% and 46%, and sham arm response rates were reported in six studies and ranged between 0% and 80%. The response rate range among somatic arms was considered inflated though since four of the six somatic studies included only 10–25 patients per arm^{56,58,59,65}. Hence, the addition of one patient with a response would greatly impact the response rate. Remission rates were higher in placebo arms than in sham arms: remission rates in placebo arms, reported in five studies, ranged from 13.5%⁵³ to 23.8%³⁷ whereas remission rates across sham arms, reported in four studies, ranged from 0% to 8.3%. However, differences in response and remission rates between placebo and sham arms were not considered clinically relevant and pooling of these arms was considered feasible. A scenario analysis was performed excluding somatic interventions to determine the validity of the similarity assumption.

The following additional assumptions were made when constructing the global network: it was considered feasible to pool lithium augmentation schedules from four different studies^{40,46,49,67}, ketamine treatment schedules (one 0.5 mg/kg infusion twice or thrice per week) from 1 study⁵¹ and different intensities of rTMS treatment (80–120%^{54–65,58–60,62–66}) from 11 different studies. The validity of the latter assumption was tested in a scenario analysis by separating rTMS intensities. It was not considered feasible to pool different dosages of quetiapine XR, so quetiapine XR 150 mg and quetiapine XR 300 mg are represented separately in the network. Also, pooling of the quetiapine arms of Doree *et al.*⁴⁰ and Bauer *et al.*³⁷ was also not considered feasible due to different induction and maximum treatment dosages.

Network meta-analysis results

Mean difference in change from baseline in Montgomery-Åsberg depression rating scale score at 2 weeks

For the primary outcome of the analysis, there were 16 nodes and 17 direct treatment comparisons in the network plot (Figure 1) for CFB at 2 weeks. A random effects model was employed and compared with placebo/sham. Intravenous ketamine, risperidone augmentation and rTMS (80–120%) were significantly more efficacious showing mean differences in CFB of -14 , -10.3 and -4.2 , respectively. Quetiapine augmentation schemes of 150, 300 and 800 mg, and ECT seemed to demonstrate better efficacy compared to placebo/sham showing mean differences in CFB of -2.4 , -2.9 , -6.7 and -4.6 , respectively, although the 95% CrIs of

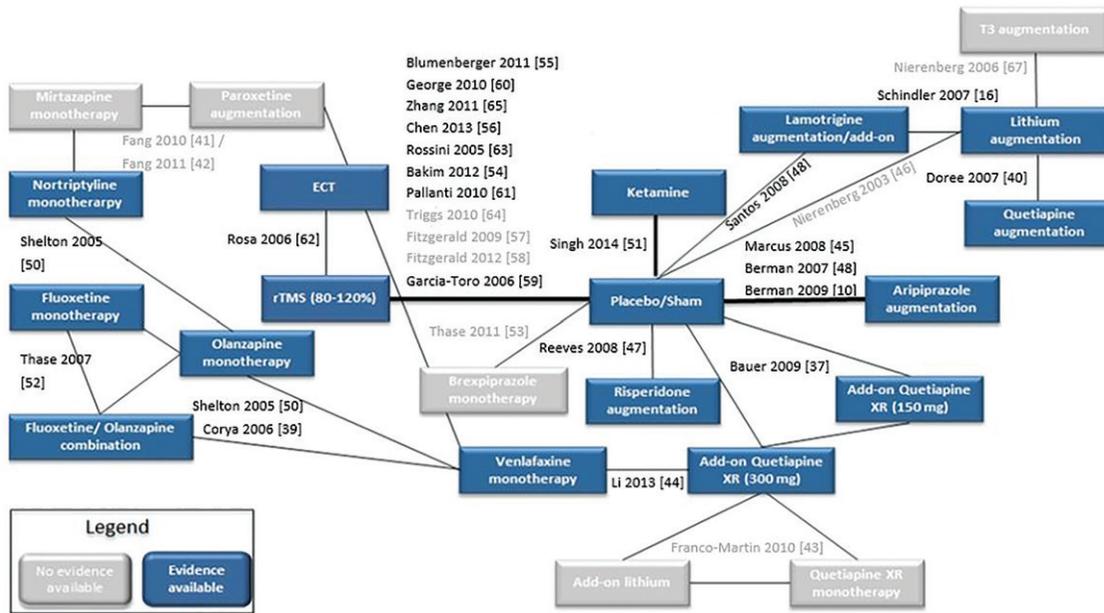


Figure 1. Network – disease severity change from baseline at 2 weeks.

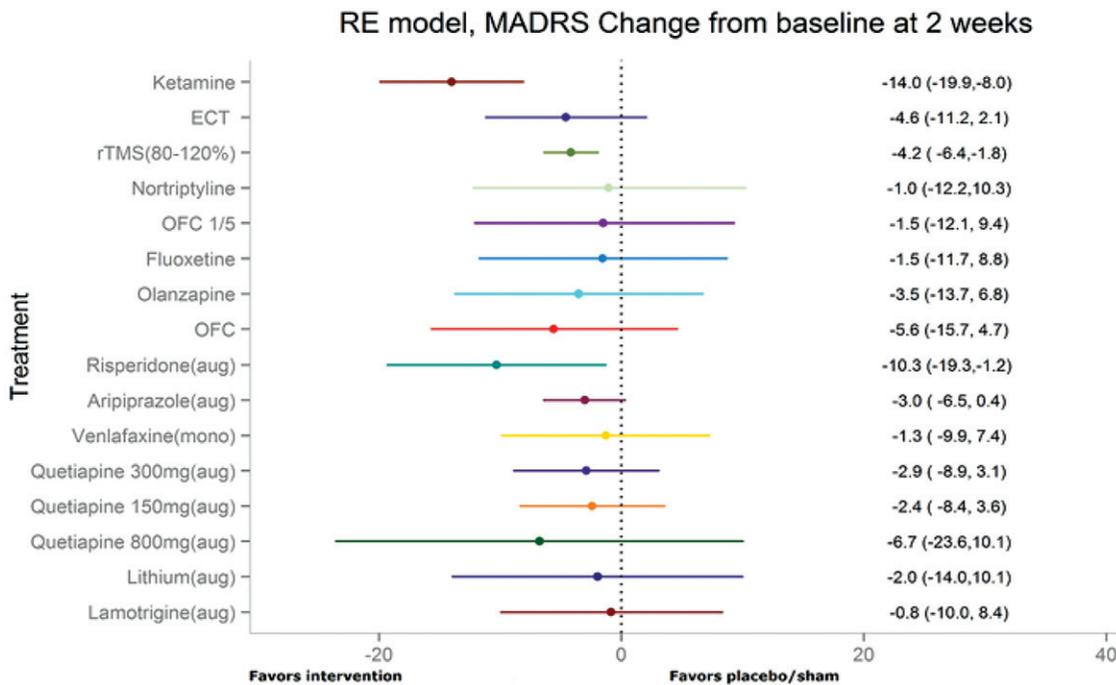


Figure 2. Network meta-analysis results – mean differences in change from baseline at 2 weeks.

RE: random effects; MADRS: Montgomery-Åsberg Depression Rating Scale; ECT: Electroconvulsive therapy; rTMS: repetitive Transcranial magnetic stimulation; OFC: Olanzapine/Fluoxetine combination; mono: monotherapy; aug: augmentation.

the treatment effect estimates included 0. Nortriptyline monotherapy, OFC, fluoxetine monotherapy, olanzapine monotherapy, venlafaxine monotherapy and lamotrigine augmentation were found comparable (mean difference close to 0) to placebo/sham arms (Figure 2). Ranking outcomes summarizing the median rank, the probability of being best and SUCRA values for this analysis are presented in Supplementary eTable 9. Intravenous ketamine ranked first, followed by risperidone augmentation and quetiapine augmentation of 800 mg/day.

Mean difference in change from baseline in Montgomery-Åsberg depression rating scale score at 4, 6 and 8 weeks

The network meta-analysis was repeated for CFB at 4, 6 and 8 weeks (networks not shown, available upon request). At 4 weeks, only rTMS (80–120%) showed a statistically significant mean difference in CFB compared with placebo (CFB difference -5.8, 95% CrI -8.6, -2.9). Risperidone augmentation and quetiapine augmentation 800 mg seemed better than placebo yet the 95% CrIs included 0 (-8.1 [-18.9, 2.6], -10.6 [-29.4, 8.2],

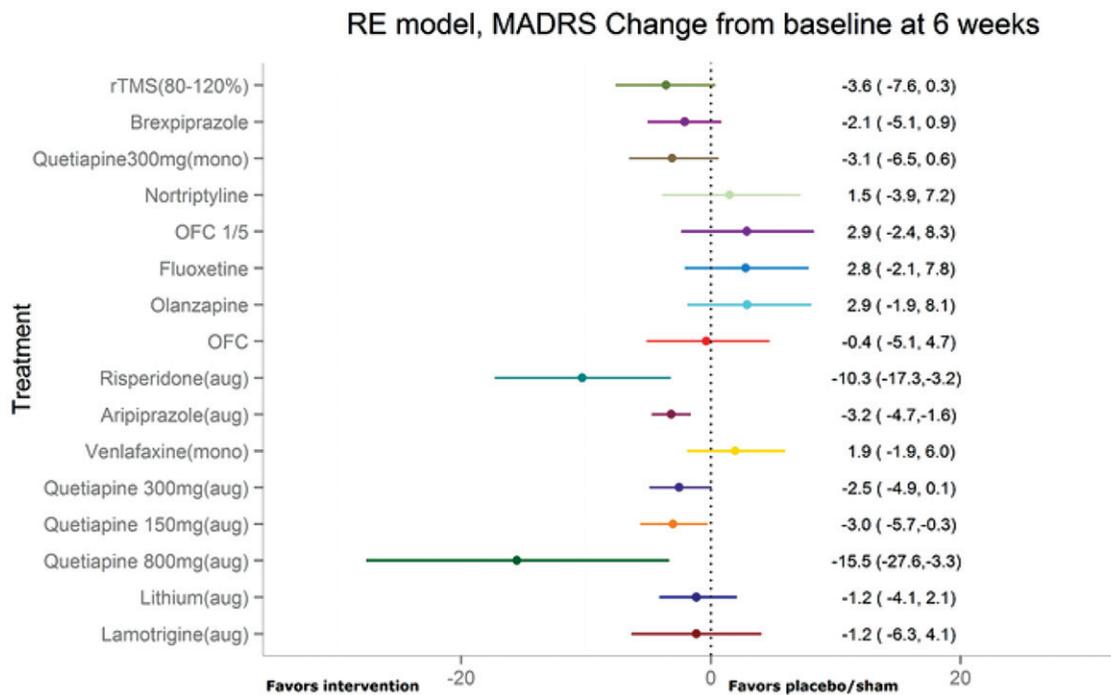


Figure 3. Network meta-analysis results – mean differences in change from baseline at 6 weeks.

RE: random effects; MADRS: Montgomery–Åsberg Depression Rating Scale; rTMS: repetitive Transcranial magnetic stimulation; OFC: Olanzapine/Fluoxetine combination; mono: monotherapy; aug: augmentation.

Table 1. Network meta-analysis results – response rate at 2 weeks.

Intervention	Comparator					
	Placebo/Sham		Aripiprazole (aug)		rTMS (80–120%)	
Aripiprazole (aug)						
Estimate of treatment effect		2.6				
95% CrI	1.8		3.8			
P (better)		>99%				
rTMS (80–120%)						
Estimate of treatment effect		2.9		1.1		
95% CrI	1.5		6.0	0.5	2.5	
P (better)		>99%		58%		
Ketamine						
Estimate of treatment effect		13.7		5.2		4.8
95% CrI	3.8		69.1	1.4	27.5	27.5
P (better)		>99%		>99%		98%

aug:augmentation; rTMS: repetitive Transcranial magnetic stimulation.

respectively). In the network of studies with CFB data available at 6 weeks, two additional treatments were included, brexpiprazole and quetiapine 300 mg monotherapy, that reported data only at this timepoint. At 6 weeks from baseline, compared to placebo/sham, quetiapine augmentation schemes of 150, 300 and 800 mg, risperidone augmentation and aripiprazole augmentation reduced depressive symptoms by -2.4 , -2.6 and -10.6 , respectively (Figure 3). CFB at 8 weeks after baseline was only reported in pharmacological intervention studies. At this timepoint, quetiapine augmentation 800 mg and risperidone augmentation showed non-significant mean differences in CFB of -11.5 and -7.6 , respectively compared to placebo.

Response rate at 2 weeks

The network of studies for treatment response rates at 2 weeks consisted of four nodes/treatments and three pairwise direct comparisons (Supplementary eFigure 7).

For studies that did not report response rates, data was imputed using the imputation method described in the *Methods* section. All three active treatments were more efficacious compared to placebo/sham (Table 1). Compared to placebo/sham, the percentage of responders was 14-fold higher for ketamine and approximately three-fold higher for both aripiprazole augmentation and rTMS (80–120%). Among the active treatments, intravenous ketamine showed a five-fold higher percentage of responders compared to aripiprazole augmentation and rTMS (80–120%).

Response rate at 4, 6 and 8 weeks

At 4 weeks after baseline, response rate data was available for aripiprazole augmentation, rTMS (80–120%) and ECT. Compared to placebo/sham, response rates were, respectively, two-fold, 2.7-fold and four-fold increased for aripiprazole augmentation, ECT and rTMS (80–120%). At 6 weeks after baseline, a larger, interlinked network of 17 active

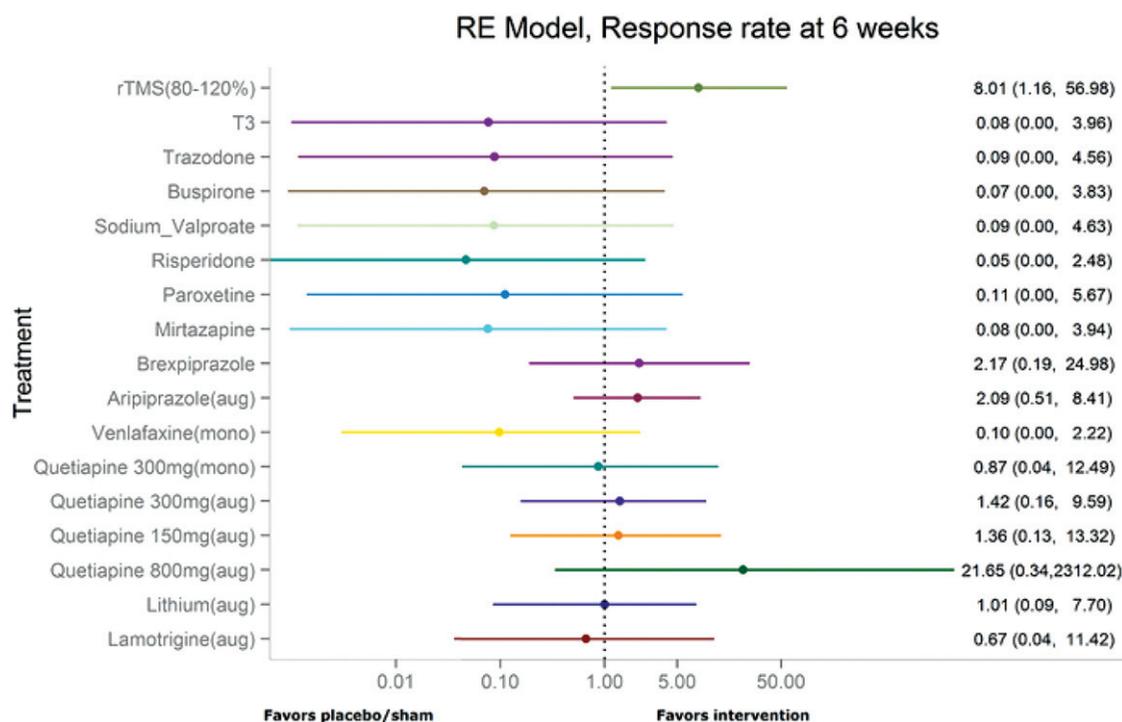


Figure 4. Network meta-analysis results – response rate at 6 weeks.

RE: random effects; rTMS: repetitive Transcranial magnetic stimulation; T3: Triiodothyronine; mono: monotherapy; aug: augmentation.

treatments could be formed with higher response rates for rTMS (80–120%) compared to placebo/sham (OR 8.01, 95% CrI 1.16–56.98). Augmentation with brexpiprazole, aripiprazole, quetiapine XR 150 mg/day, XR 300 mg/day and 800 mg/day seemed more efficacious compared to placebo/sham (ORs of 2.17, 2.09, 1.36, 1.4 and 21.65), yet the 95% CrIs included 1 (Figure 4). At 8 weeks, only pharmacological intervention studies reported data on response rates. Augmentation with quetiapine 800 mg/day was better compared to all competing interventions (placebo, OR 9.74; lamotrigine augmentation, OR 12.67; and lithium augmentation, OR 20.53).

Remission rate at 2 weeks

At 2 weeks, only three pharmacological intervention studies and three somatic intervention studies reported data on the proportion of patients that remitted. Augmentation with aripiprazole and rTMS (80–120%) were found to be more efficacious than placebo/sham (Supplementary eTable 10). No differences were found in remission rates between augmentation with aripiprazole and rTMS (80–120%).

Remission rate at 4, 6 and 8 weeks

At 4 weeks after baseline, aripiprazole augmentation, rTMS (80–120%) and ECT treatments were more efficacious than placebo/sham (OR: 2.55, 9.51 and 24.43, respectively). At 6 weeks after baseline, rTMS (80–120%) showed the highest remission rates and ranked first among all competing pharmacological interventions (OR 8.58 95% CrI 1.15, 112.55). Augmentation with brexpiprazole, aripiprazole, quetiapine XR 150 mg/day, XR 300 mg/day and 800 mg/day were efficacious compared to

placebo/sham, yet the 95% CrI included 1 (Supplementary eFigure 8). At 8 weeks after baseline, only two active pharmacological treatments reported remission rates compared to placebo. Augmentation with quetiapine 800 mg/day was better compared to all competing interventions (lamotrigine augmentation, OR 0.21; and lithium augmentation, OR 12.00).

For all presented efficacy analyses, scenario (sensitivity) analyses separating pharmacological from somatic interventions were performed. The results of the scenario analyses for the pharmacological interventions were in line with the joint analysis. The results of all scenario analyses are presented in Supplementary material.

Withdrawals due to adverse events (at 6 weeks)

The network formed for withdrawals due to adverse events consisted of four active treatments: three pharmacological treatments, augmentation with aripiprazole, quetiapine (150 mg/day and 300 mg/day), and lamotrigine, and one somatic treatment, rTMS (80–120%). Compared to placebo/sham, lamotrigine augmentation had a similar tolerability profile (OR 1.00, 95% CrI 0.09–10.67). The other active treatments showed higher withdrawal rates due to adverse events compared to placebo/sham: augmentation with aripiprazole showed an approximately three-fold increase, quetiapine 150 mg a four-fold, quetiapine 300 mg a two-fold and rTMS (80–120%) a four-fold increase (Supplementary eFigure 9).

Discussion

A network meta-analysis was performed based on 31 RCTs identified through a systematic literature review including

5515 adult patients with unipolar TRD assigned to 13 pharmacological interventions and two somatic interventions. The main finding is that ketamine demonstrated superior efficacy, i.e. a faster reduction in depression severity and a higher response rate (defined as a $\geq 50\%$ reduction in depression severity), over the competing pharmacological and somatic interventions at 2 weeks after the start of treatment. Ketamine treatment demonstrated rapid antidepressant effects and was ranked first among all 12 pharmacological and two somatic interventions. The findings for later timepoints of 4, 6 and 8 weeks, for which no ketamine data were available, favored quetiapine 800 mg/day, risperidone augmentation and rTMS treatment of pooled intensities.

The findings of this study are aligned with results from published systematic reviews and meta-analyses of ketamine for the rapid treatment of major depressive disorders. McGirr *et al.*⁶⁸ performed a meta-analysis of randomized placebo-controlled trials of ketamine investigating clinical response and remission at 24 hours, 3 days and 7 days post-treatment. McGirr *et al.* reported a pooled OR of 4.87 of ketamine versus placebo, after 7 days, regarding response rate and 4.00 regarding remission rate defined as HAM-D score < 7 or MADRS score < 10 . Ketamine was administered intravenously in all but one study, which employed intranasal ketamine. Intravenous infusion protocols most commonly used 0.5 mg/kg over 40 minutes. At earlier timepoints, 24 hours, ketamine resulted in a nine-fold higher response rate versus placebo (OR: 9.10, 95% CI: 4.28 to 19.34).

Given that the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial revealed a ceiling effect in treating patients despite augmentation, switching and combined strategies, ketamine should be further investigated in longer-term RCTs. We acknowledge that there are ongoing trials (SUSTAIN 1–3^{69–71}) investigating the long-term efficacy and safety of intranasal esketamine in TRD.

The results regarding the efficacy of atypical antipsychotics are aligned with the findings of Zhou *et al.*²¹. Risperidone augmentation and a high dose of quetiapine augmentation (800 mg/day) appeared to be efficacious antipsychotic agents for the timepoint analysis of 4, 6 and 8 weeks. Aripiprazole augmentation and OFC were not found as efficacious in alleviating depressive symptoms. From the above mentioned agents, only OFC is licensed for the treatment of TRD.

Regarding results of somatic interventions, ECT was found to perform significantly better compared to placebo/sham and comparable to pooled intensities of rTMS for the two timepoints (2 and 4 weeks) where data was available. The evidence base included only one ECT trial hence wide credible intervals were associated with its estimates. Pooled intensities of rTMS were found to be significantly more efficacious than placebo/sham throughout all analyses and numerically better than most pharmacological comparators with the exception of ketamine at 2 weeks.

This study is the first study to compare the relative efficacy of pharmacological and somatic interventions within one network meta-analysis. These interventions were compared using a common comparator arm of pooled placebo

and sham arms from, respectively, pharmacological and somatic intervention studies. The feasibility steps revealed no relevant differences between the placebo arms of pharmacological intervention studies and the sham arms of somatic intervention studies. In addition, a scenario analysis excluding somatic interventions to determine the validity of the similarity assumption did not show a large change in results.

There are a number of strengths of the current network meta-analysis. First of all, a comprehensive and robust search strategy was used to identify all relevant pharmacological and somatic interventions for treatment resistant depression. An additional strength of this analysis was the rigorous definition of TRD as at least two treatment failures; most published systematic reviews and (network) meta-analyses report and pool studies irrespectively of the number of treatment failures, using as TRD definition an inadequate response to initial treatment. The overall quality of identified and included studies in the analysis was rated as good despite the majority of pharmacological and somatic interventions studies not adequately reporting the randomization and allocation concealment schemes.

A general limitation of network meta-analyses is that the indirect comparisons made may be influenced by potential biases and uncertainties introducing heterogeneity and inconsistent outcomes between studies. For the majority of the analyses, random effects models could be applied to account for potential heterogeneity across trials. In addition, all inconsistency checks performed did not reveal inconsistency in the treatment effects estimates. Also, by using a more rigorous TRD definition, we aimed to reduce heterogeneity among studies.

The evidence base of the current network meta-analysis had several limitations that could influence the results. First of all, for efficiency reasons the systematic literature search was restricted to articles published between January 2003 and September 2014. However, although publications might have been missed, older publications are likely to have included different regimens and different study populations that would not have adhered to our rigorous TRD definition. Also, placebo responses across antidepressant trials have been shown to increase in the past two decades⁷². In the current evidence base not all studies reported change from baseline according to the MADRS depression rating scale. Although the applied mapping technique to convert HAM-D₁₇ scores to MADRS scores facilitated valid comparisons, a different mapping formula would have resulted in different scores. In addition, many somatic intervention studies showed small sample sizes ranging from 10 to 99 patients per arm, with a poor evidence base for ECT consisting only of one study including only 21 TRD patients.

Conclusions

This Bayesian network meta-analysis is the first to compare the relative efficacy of pharmacological and somatic interventions for TRD. Based on this analysis, ketamine has a rapid and robust antidepressant effect as demonstrated by superior short-term results at 2 weeks. However further studies are

needed to evaluate its long-term antidepressant efficacy and safety. Regarding efficacy results at later timepoints, there was no clear distinction among the investigated treatments except high dose quetiapine augmentation (800 mg/day) and risperidone augmentation (0.25–2 mg/day) which showed superior efficacy compared to the competing interventions at 4, 6 and 8 weeks analysis. This analysis revealed scarcity of long-term data (i.e. data on sustained remission) that would allow a comparative long-term efficacy assessment.

Transparency

Declaration of funding

This study was conducted by Mapi on behalf of Janssen Pharmaceutica NV who funded the study and the writing of this manuscript.

Author contributions: All authors were involved in defining the scope of the project. K.P., C.V. and A.K. conducted the systematic review and analyzed the data. C.K. and N.K. contributed material. K.P., C.V. and A.K. wrote the manuscript and C.D. and N.K. were involved in critical revision and final approval of the manuscript.

Christina Donatti^b

Nicole Kubitz^c

Declaration of financial/other relationships

K.P., C.V. and A.K. have disclosed that they are employees of Mapi and served as paid consultants to Janssen Pharmaceutica NV during the conduct of this study and the preparation of this manuscript. CD is an employee of Janssen UK, High Wycombe, UK and NK an employee of Janssen-Cilag GmbH, Neuss, Germany.

CMRO peer reviewer 1 has disclosed that he has been a recipient of grants or research funding from Servier and Eli Lilly; has been a consultant or advisor to Servier, Eli Lilly, AstraZeneca, sanofi aventis, Janssen-Cilag and Lundbeck; and has been a member of the speakers bureau for Bristol Myers Squibb, Eli Lilly, Lundbeck and AstraZeneca. CMRO peer reviewer 2 has no relevant financial or other relationships to disclose.

Acknowledgements

The authors would like to thank the reviewers for their valuable comments.

References

- Gruenberg AM, Goldstein RD, Pincus HA. Classification of Depression: Research and Diagnostic Criteria: DSM-IV and ICD-10. *Biology of Depression: From Novel Insights to Therapeutic Strategies*. Wiley-VCH Verlag GmbH, 2008
- Bromet E, Andrade LH, Hwang I, et al. Cross-national epidemiology of DSM-IV major depressive episode. *BMC Med* 2011;9:90
- Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380:2163-96
- O'Leary OFDT, Cryan JF. Faster, better, stronger: towards new antidepressant therapeutic strategies. *Eur J Pharmacol* 2015;32-50
- Shelton RCOO, Heinloth AN, Corya SA. Therapeutic options for treatment-resistant depression. *CNS Drugs* 2010;24:131-61
- Rush AJTM, Wisniewski SR, Nierenberg AA, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry* 2006;163:1905-17
- Lehmann HE. Therapy-resistant depressions – a clinical classification. *Pharmakopsychiatri* 1974;7:156-63
- Sartorius N. Description and classification of depressive disorder. Contributions for the definition of therapy-resistance and of therapy-resistant depressions. *Pharmakopsychiatri* 1974;7:76-9
- Thase MERA. When at first you don't succeed: sequential strategies for antidepressant nonresponders. *J Clin Psychiatry* 1997;58(Suppl 13):23-9
- Berman RM, Fava M, Thase ME, et al. Aripiprazole augmentation in major depressive disorder: a double-blind, placebo-controlled study in patients with inadequate response to antidepressants. *CNS Spectrums* 2009;14:197-206
- Souery DPG, Trivedi MH. Treatment-resistant depression. *J Clin Psychiatry* 2006;67(Suppl 6):16-22
- Symbyax [package insert]. Indianapolis, IN: Eli Lilly and Company, 2009
- Cleare A, Pariante CM, Young AH. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2008 British Association for Psychopharmacology guidelines. *J Psychopharmacol* 2015;29:459
- Papakostas GISR, Smith J, Fava M. Augmentation of antidepressants with atypical antipsychotic medications for treatment-resistant major depressive disorder: a meta-analysis. *J Clin Psychiatry* 2007;68:826-31
- Nelson JCPG. Atypical antipsychotic augmentation in major depressive disorder: a meta-analysis of placebo-controlled randomized trials. *Am J Psych* 2009;166:980-91
- Spielmanns GIBM, Linardatos E, Rosenlicht NZ, et al. Adjunctive atypical antipsychotic treatment for major depressive disorder: a meta-analysis of depression, quality of life, and safety outcomes. *PLoS Med* 2013;10:e1001403
- Papakostas GINJ, Kasper S, Möller HJ. A meta-analysis of clinical trials comparing reboxetine, a norepinephrine reuptake inhibitor, with selective serotonin reuptake inhibitors for the treatment of major depressive disorder. *Eur Neuropsychopharmacol* 2008;18:122-12
- Bauer MDS. Lithium augmentation in treatment-resistant depression: meta-analysis of placebo-controlled studies. *J Clin Psychopharmacol* 1999;19:421-34
- Gaynes BNLS, Lux L, Gartlehner G, Hansen RA. Repetitive transcranial magnetic stimulation for treatment-resistant depression: a systematic review and meta-analysis. *J Clin Psychiatry* 2014;75:477-89
- Zhou XRA, Qin B, Del Giovane C, et al. Comparative efficacy, acceptability and tolerability of augmentation agents in treatment-resistant depression: systematic review and network meta-analysis. *J Clin Psychiatry* 2014;76:487-98
- Zhou XKG, Qin B, Ravindran A. Atypical antipsychotics for treatment-resistant depression: a systematic review and network meta-analysis. *Int J Neuropsychopharmacol* 2015;18:1-10
- Kirsch IMJ. Clinical trials and the response rate illusion. *Contemp Clin Trials* 2007;28:348-51
- Zimmerman MPMCI. Derivation of a definition of remission on the Montgomery-Asberg depression rating scale corresponding to the definition of remission on the Hamilton rating scale for depression. *J Psychiatric Res* 2004;38:577-82
- Furukawa TA, Cipriani A, Barbui C, et al. Imputing response rates from means and standard deviations in meta-analysis. *Int Clin Psychopharmacol* 2005;20:49-52
- Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373:746-58
- The Cochrane Collaboration's for assessing risk of bias in randomized trials. *BMJ* 2011;343:d5928
- Lu G, Ades A. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105-24
- Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279-301
- Hoaglin DC, Hawkins N, Jansen JP, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the

- depression: a randomized, boucle-blind, sham-controlled study. *Shanghai Arch Psychiatry* 2011;17-24
66. Zheng H, Zhang L, Li L, et al. High-frequency rTMS treatment increases left prefrontal myo-inositol in young patients with treatment-resistant depression. *Progr Neuro Psychopharmacol Biol Psychiatry* 2010;34:1189-95
 67. Nierenberg AA, Fava M, Trivedi MH, et al. A comparison of lithium and T(3) augmentation following two failed medication treatments for depression: a STAR*D report. *Am J Psychiatry* 2006;163:1519-30; quiz 1665
 68. McGirr A, Berlim MT, Bond DJ, et al. A systematic review and meta-analysis of randomized, double-blind, placebo-controlled trials of ketamine in the rapid treatment of major depressive disorders. *Psychol Med* 2014;45:693-704
 69. Janssen Research & Development. A Long-term, Safety and Efficacy Study of Intranasal Esketamine in Treatment-resistant Depression (SUSTAIN-2). [clinicaltrials.gov : NCT02497287](https://clinicaltrials.gov/ct2/show/study/NCT02497287) October 2016
 70. Janssen Research & Development. A Long-term Safety Study of Intranasal Esketamine in Treatment-resistant Depression (SUSTAIN-3). [clinicaltrials.gov : NCT02782104](https://clinicaltrials.gov/ct2/show/study/NCT02782104) 2016
 71. Janssen Research & Development. A Study of Intranasal Esketamine Plus an Oral Antidepressant for Relapse Prevention in Adult Participants With Treatment-resistant Depression (SUSTAIN-1). [clinicaltrials.gov : NCT02493868](https://clinicaltrials.gov/ct2/show/study/NCT02493868) 2016
 72. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies with major depression: variable, substantial and growing. *JAMA* 2002;287:1840-7