Comparative efficacy and acceptability of pharmacological treatments for post-traumatic stress disorder in adults: a network meta-analysis

Author for correspondence:
Andrea Cipriani, E-mail: andrea.cipriani@psych.ox.ac.uk

Andrea Cipriani[1], Taryn Williams[2], Adriani Nikolakopoulou[3], Georgia Salanti[3], Anna Chaimani[4], Jonathan Ipser[2], Phil J. Cowen[1], John R. Geddes[1] and Dan J. Stein[2]

[1]Department of Psychiatry, University of Oxford, UK; [2]Department of Psychiatry and Mental Health and Medical Research Council of South Africa Unit on Anxiety & Stress Disorders, University of Cape Town, South Africa; [3]Department of Clinical Research, Institute of Social and Preventive Medicine, University of Bern, Switzerland and [4]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Greece

Abstract

Background. Guidelines about post-traumatic stress disorder (PTSD) recommend broad categories of drugs, but uncertainty remains about what pharmacological treatment to select among all available compounds.
Methods. Cochrane Central Register of Controlled Trials register, MEDLINE, PsycINFO, National PTSD Center Pilots database, PubMed, trial registries, and databases of pharmaceutical companies were searched until February 2016 for double-blind randomised trials comparing any pharmacological intervention or placebo as oral therapy in adults with PTSD. Initially, we performed standard pairwise meta-analyses using a random effects model. We then carried out a network meta-analysis. The main outcome measures were mean change on a standardised scale and all-cause dropout rate. Acute treatment was defined as 8-week follow up.
Results. Desipramine, fluoxetine, paroxetine, phenelzine, risperidone, sertraline, and venlafaxine were more effective than placebo; phenelzine was better than many other active treatments and was the only drug, which was significantly better than placebo in terms of dropouts (odds ratio 7.50, 95% CI 1.72–32.80). Mirtazapine yielded a relatively high rank for efficacy, but the respective value for acceptability was not among the best treatments. Divalproex had overall the worst ranking.
Conclusions. The efficacy and acceptability hierarchies generated by our study were robust against many sources of bias. The differences between drugs and placebo were small, with the only exception of phenelzine. Considering the small amount of available data, these results are probably not robust enough to suggest phenelzine as a drug of choice. However, findings from this review reinforce the idea that phenelzine should be prioritised in future trials in PTSD.

CAMBRIDGE
UNIVERSITY PRESS

Background

Post-traumatic stress disorder (PTSD) is a mental disorder that develops in up to a third of individuals who are exposed to extreme stressors, with significant work and social impairment, increased risk of suicide, higher medical and social costs and higher psychiatric comorbidity (Norris & Sloane, 2007; Steel *et al.* 2009; Krysinska & Lester, 2010). The estimated lifetime prevalence of PTSD among adults in the general population is around 6.8% and current (12-month) prevalence about 3.6% (Kessler *et al.* 2011). However, recent surveys of military personnel have yielded higher estimates, ranging from 6.2% for US service members who fought in Afghanistan to 12.6% for those who fought in Iraq and 15.4% in conflict-affected populations (Dohrenwend *et al.* 2006; Seal *et al.* 2007). Treatments available for PTSD span a variety of psychological and pharmacological interventions (Jonas *et al.* 2013). Numerous organizations have issued guidance for the treatment of patients with PTSD (APA, 2004; NICE, 2005; PA-CPMH, 2013), supporting trauma-focused psychological interventions as first-line treatments for PTSD. These guidelines have also recognised some benefit of pharmacological interventions; however, they have arrived at different conclusions about the value of such interventions, and have limited themselves to recommendations about broad categories of treatments. From a clinical point of view, though, it is very important to know what pharmacological intervention is more efficacious and acceptable than others among the various treatment options available for PTSD. Notwithstanding the recent publication of two systematic reviews (Watts *et al.* 2013; Hoskins *et al.* 2015), clinical uncertainty still remains about

what pharmacological treatment to select among all available compounds because traditional pair-wise meta-analyses do not allow the simultaneous integration of direct and indirect evidence (Cipriani *et al.* 2013). The aim of this paper is to assess efficacy and acceptability of different pharmacological treatments against one another, to provide a clinically useful summary of the comparative evidence that can be used to guide decisions about acute treatment of PTSD in adults.

Methods

The study protocol was drafted and made available on our institutional website (http://cebmh.warne.ox.ac.uk/cebmh/research_other_reviews.htm) (Appendix 1). Furthermore, with the publication of this paper, the overall dataset will be in the public domain for anyone who would be interested to use it.

Study eligibility criteria

We identified all double-blind randomised controlled trials (RCTs) comparing any psychotropic agent at a therapeutic dose with another psychotropic drug or placebo as oral therapy in the treatment of adults with PTSD, diagnosed according to operationalised criteria. We included both monotherapy and add-on studies, and assumed that add-on interventions had an additive effect on the treatments of interest. The dose ranges for the medication included in the network meta-analysis were defined according to the FDA, whenever possible (https://dailymed.nlm.nih.gov/dailymed/ or http://www.accessdata.fda.gov/). If information was not available, we used the British National Formulary (https://www.bnf.org/). We included all the following pharmacological interventions: amitriptyline, clomipramine, desipramine, imipramine, maprotiline, mianserin, tianeptine, trazodone, nefazodone, citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline, bupropion, desvenlafaxine, venlafaxine, duloxetine, mirtazapine, reboxetine, brofaromine, moclobemide, phenelzine (among antidepressants); haloperidol, chlorpromazine, olanzapine, risperidone, quetiapine (among antipsychotics); lithium, lamotrigine, valproate, tiagabine, topiramate (among mood stabilisers/anti-epileptic drugs); guanfacine, prazosin and selective NK1R antagonists (among the new drugs with different mechanisms of action). The synthesis comparator set consists of all the interventions listed above, their combinations and placebo (for all details about Methods, see Appendix 1).

Identification and selection of studies

We searched Cochrane Depression, Anxiety and Neurosis Group specialised register and the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, PsycINFO, the National PTSD Center Pilots database and PubMed (last update: 15 February 2016). We also searched the databases of the most important regulatory agencies worldwide and online trial registers for published and unpublished RCTs (see Appendix 2 for full details on the search strategy). No language restrictions were applied. All relevant authors and principal manufacturers were contacted to supplement the incomplete report of the original papers.

Data collection and study appraisal

Acute treatment was defined as 8-week treatment in all analyses. Mean change scores on the Clinician-Administered PTSD Scale

(CAPS) and dropout rates (treatment discontinuation) were chosen as primary outcomes to represent the most sensible and sensitive estimate of acute treatment efficacy and acceptability, respectively. As secondary analyses, we also estimated the proportion of patients who responded to treatment (according to original study authors' definition) and the proportion of patients who left the study early due to adverse events (tolerability). The risk of bias tool was used to assess study quality (Appendix 3) (Higgins & Green, 2011). Additionally, we assessed the quality of evidence contributing to each network estimate using the GRADE framework, which characterises the quality of a body of evidence on the basis of the study limitations, imprecision, inconsistency, indirectness, and publication bias for the primary outcomes (Salanti *et al.* 2014).

Data synthesis and statistical analysis

For continuous outcome, where different measures were used to assess the same outcome, standardised mean difference (SMD) (Hedges's adjusted *g*) were used (Altman & Bland, 1996; Higgins & Green, 2011). Dichotomous outcomes were analysed by calculating the odds ratio (OR).

Assessment of clinical and methodological heterogeneity and transitivity

To evaluate the presence of heterogeneity we generated descriptive statistics for trial and study population characteristics across all eligible trials (Furukawa *et al.* 2006). If the distributions were balanced across comparisons, we concluded against evidence of intransitivity (Salanti, 2012; Jansen & Naci, 2013).

Data synthesis

Initially, we performed standard pairwise meta-analyses using a random effects model (DerSimonian & Lair, 1986) in STATA (StataCorp, 2015). We then carried out a network meta-analysis (NMA) to synthesise the available evidence from the entire network of trials by integrating direct and indirect estimates and using the methodology of multivariate meta-analysis (different treatment comparisons were treated as different outcomes, assuming a common heterogeneity parameter) (White *et al.* 2012). We also estimated the ranking probabilities using the surface under the cumulative ranking curve (SUCRA) (Salanti *et al.* 2011) and mean ranks. To perform NMA we used the *mvmeta* command in Stata (White, 2011) while checking the model assumptions and presenting the results we used a suite of previously published Stata commands (Chaimani *et al.* 2013*a*).

Statistical heterogeneity and inconsistency

To measure the percentage of variability that cannot be attributed to random error, we assessed statistically the presence of heterogeneity in all pairwise comparisons by calculating the heterogeneity standard deviation parameter (τ) and within each comparison using the I-squared (I^2) statistic (Higgins *et al.* 2003) and its 95% confidence interval (CI). For all outcomes τ was compared with its empirical distribution (Turner *et al.* 2012; Rhodes *et al.* 2015). We also estimated a total I^2 value for heterogeneity in the network (Jackson *et al.* 2014). We evaluated the presence of inconsistency locally (*local tests*) using the loop-specific approach (Song *et al.* 2003; Chaimani *et al.* 2013*b*; Veroniki *et al.* 2013). To check

the assumption of consistency in the entire network (*global test*) we used the *design-by-treatment model* (White, 2011; Higgins *et al.* 2012). To distinguish between inconsistency and heterogeneity, we employed the I^2 for inconsistency and measured the percentage of variability that cannot be attributed to random error or heterogeneity (Turner *et al.* 2012). In case of important heterogeneity or/and inconsistency, we carried out meta-regression or subgroup analyses by using the following effect modifiers: (a) variance of the observed relative effects; (b) baseline severity; (c) publication year and (d) sponsorship.

Sensitivity analysis

To check the robustness of study findings, we carried out a sensitivity analysis excluding studies rated as at high or unclear risk of bias (Wood *et al.* 2008).

Results

Description of the evidence base

The systematic review included overall 51 studies with 6189 patients (see Appendices 4 and 5), randomised to 25 different interventions (Fig. 1). The mean duration of primary studies was 10.3 weeks (s.d. 2.1) and the mean age of trial participants was 42.9 years (s.d. 5.9). The great majority of studies were 2-arm (45 studies), but the search retrieved also six 3-arm studies.

Thirty-six studies (71.0%) used CAPS (or CAPS-2) as the rating scale for assessing efficacy, with a mean baseline severity of 74.1 (s.d. 16.7). Only 11 trials (22.0%) had an add-on design and overall 36 studies (71.0%) were sponsored by pharmaceutical industry. In terms of study quality, the reports often did not provide full details about randomisation procedures (28 out of 51, 54.9%) but the reporting about allocation concealment was much worse, as only 19 studies out of 51 (37.2%) were judged at low risk of bias (Appendix 3). The networks of eligible comparisons by outcome are shown in Fig. 2 and Appendix 6. The results of the direct pairwise comparisons for all outcomes are reported in Appendix 7. There was considerable variability in the placebo response rates among studies, which varied between 3% and 60% and resulted in a $\tau = 0.1$. However, there was no evidence of a time trend in placebo response (p value 0.93). Heterogeneity was not low only for the dichotomous response rate and results did not materially change when we assumed comparison-specific heterogeneity parameters.

Network meta-analysis results

Table 1 and Fig. 3 show the relative treatment effects from NMA for the two primary outcomes (change in symptom's severity and all-cause dropouts). In terms of efficacy, phenelzine, desipramine, paroxetine, venlafaxine, fluoxetine, risperidone, and sertraline were more effective than placebo (Fig. 3). Phenelzine appeared to be statistically significantly more effective than nearly half of
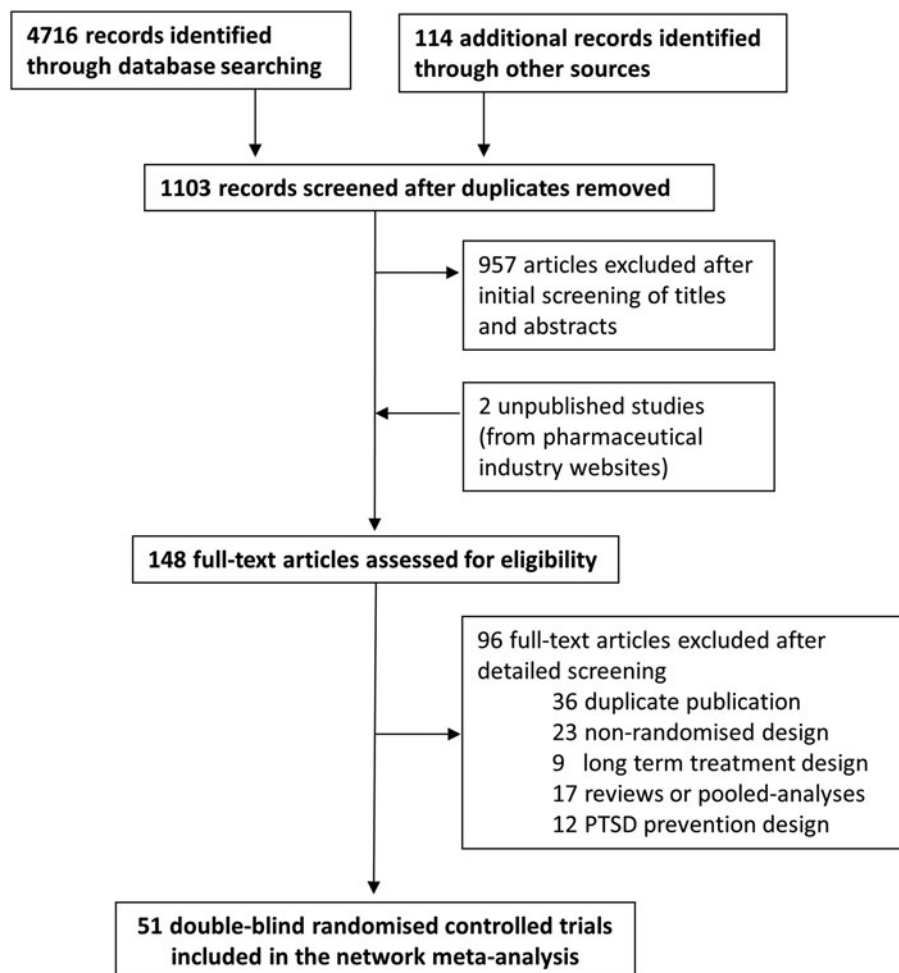

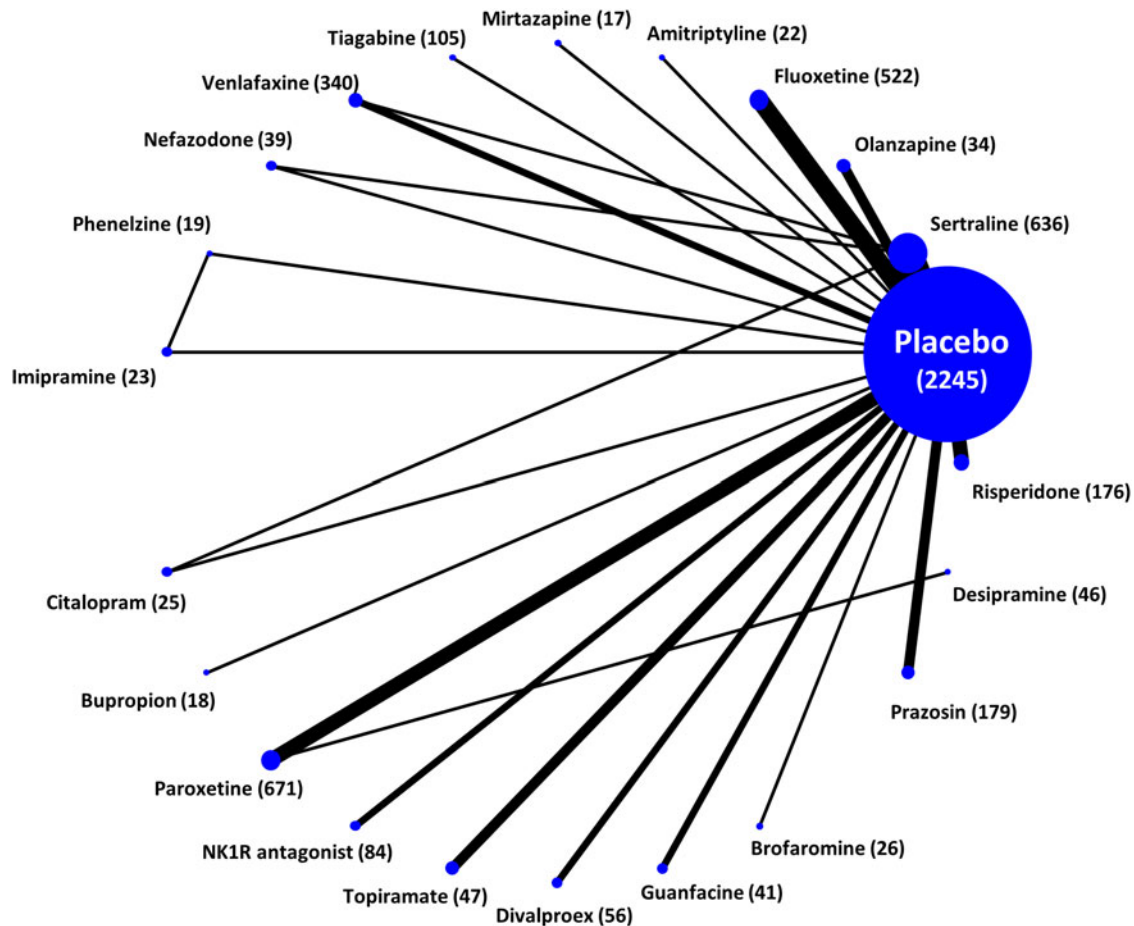Fig. 1. PRISMA flow diagram. Included and excluded studies.

Fig. 2. Network plot for the primary efficacy outcome. Nodes and edges are weighted according to the number of studies involved in each treatment comparison.

the other active treatments included in the network (9 out of 21), while both citalopram and divalproex were statistically significantly less efficacious than phenelzine, desipramine, and paroxetine (citalopram was also inferior to mirtazapine, olanzapine, venlafaxine, and fluoxetine) (Table 1). In terms of dropout rate, phenelzine was the only drug which was significantly better than placebo (OR 7.50, 95% CI 1.72–32.80) and most of the other active treatments (Table 1). No other statistically significant differences were found between the rest of the treatments and placebo (Fig. 1). Notwithstanding fewer events, response rate data did match with our efficacy findings using continuous data (Appendix 8); in terms of dropouts due to adverse events, no drug was statistically significantly better tolerated than placebo, but sertraline and paroxetine performed worse than placebo (OR 0.59, 95% CI 0.38–0.91 and 0.63, 95% CI 0.42–0.94, respectively), and brofaromine better than paroxetine (OR 2.78, 95% CI 1.04–7.42), topiramate (OR 4.77, 95% CI 1.04–21.75), and sertraline (OR 2.99, 95% CI 1.10–8.09) (Appendix 8).

We were not able to assess statistically the transitivity assumption by comparing the distributions of potential effect modifiers across comparisons as most comparisons were informed by a small number of studies. We also evaluated inconsistency in each loop of available evidence (Appendix 9). Although we did not find any statistical evidence of inconsistency assuming either network-specific or loop-specific heterogeneity, there are a few closed loops per outcome and few studies involved, so we cannot exclude the possibility of potentially important inconsistency to

be present. The design × treatment interaction model did not provide evidence that any of the networks present as a whole, statistically significant inconsistency (Appendix 10); again the small number of closed loops and studies limits the power of these tests. The I^2 for both heterogeneity and inconsistency in NMA were calculated to distinguish between these two sources of variability. For efficacy as a continuous outcome, I^2 was 22.4% for heterogeneity and 22.5% for inconsistency; for acceptability, I^2 was 0% for heterogeneity and 8% for inconsistency. The percentage of variability attributed to heterogeneity was 43.2% for the dichotomous efficacy outcome (we could not derive the respective measure for inconsistency as inconsistency could not be assessed for this outcome – see above). For dropout rate due to adverse events, I^2 was 0% for heterogeneity and 13% for inconsistency. Overall, all these measures were in general low but it should be acknowledged that the uncertainty around them might be high (see below).

For a continuous outcome in a 'drug versus placebo' comparison in mental health the suggested median value for τ is 0.22 and for a 'drug versus drug' comparison it is 0.20. Both values are greater than our estimation of heterogeneity for the efficacy as a continuous outcome using the restricted maximum likelihood (REML) estimator ($\tau = 0.1$). For a binary mental health outcome the suggested median values for τ are 0.35 and 0.31 for a 'drug versus placebo' and a 'drug versus drug' comparison, respectively. Our estimation of τ for the dichotomous efficacy outcome was 0.4, whereas the values of τ were 0 for all the remaining dichotomous outcomes, dropouts due to any cause and dropouts due to adverse events.

Table 1. NMA results for the two primary outcomes

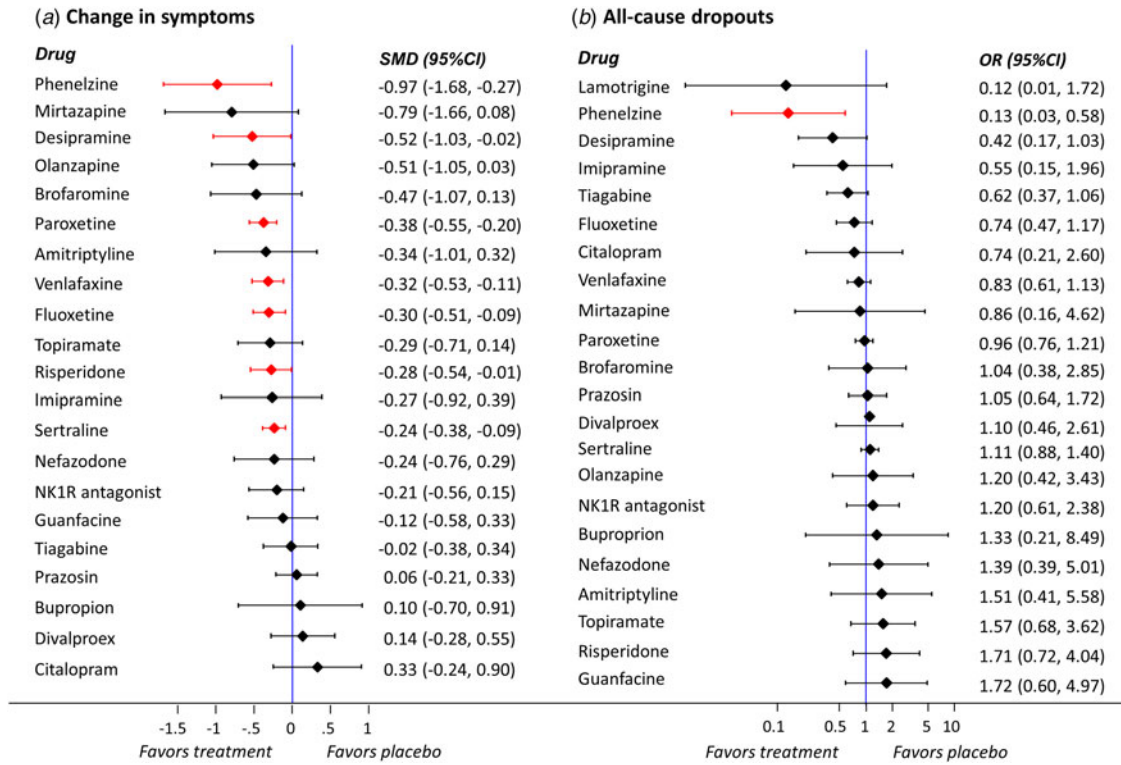| PHE	MIR	DES	OLA	BRO	PAR	VEN	AMI	FLX	IMI	TPM	RIS	NEF	SER	NK1R	GUA	TGB	BUP	PLB	PRZ	DVP	CIT	LAM	
PHE	0.19 (-0.93, 1.31)	0.45 (-0.42, 1.32)	0.46 (-0.43, 1.43)	0.50 (-0.42, 1.43)	0.60 (-0.13, 1.32)	0.66 (-0.08, 1.39)	0.63 (-0.34, 1.60)	0.67 (-0.06, 1.41)	**0.71 (0.05, 1.37)**	0.69 (-0.14, 1.51)	0.70 (-0.06, 1.45)	**0.74 (-0.14, 1.62)**	**0.74 (0.02, 1.46)**	0.77 (-0.02, 1.56)	**0.85 (0.01, 1.69)**	**0.95 (0.16, 1.75)**	**1.08 (0.06, 2.15)**	**0.97 (0.27, 1.68)**	**1.03 (0.28, 1.79)**	**1.11 (0.29, 1.93)**	**1.30 (0.39, 2.21)**	-
6.43 (0.69, 60.31)	MIR	0.27 (-0.74, 1.27)	0.27 (-0.76, 1.31)	0.32 (-0.74, 1.38)	0.41 (-0.48, 1.30)	0.47 (-0.43, 1.37)	0.45 (-0.66, 1.55)	0.49 (-0.41, 1.39)	0.52 (-0.57, 1.61)	0.50 (-0.48, 1.47)	0.51 (-0.40, 1.43)	0.55 (-0.47, 1.57)	0.55 (-0.33, 1.44)	0.58 (-0.36, 1.53)	0.66 (-0.32, 1.65)	0.77 (-0.17, 1.71)	0.89 (-0.30, 2.08)	0.79 (-0.09, 1.66)	0.84 (-0.07, 1.76)	0.92 (-0.05, 1.89)	**1.12 (0.07, 2.16)**	-
3.17 (0.57, 17.77)	0.27 (0.07, 3.31)	DES	0.05 (-0.72, 0.83)	0.01 (-0.73, 0.75)	0.14 (-0.33, 0.61)	0.20 (-0.34, 0.74)	0.18 (-0.65, 1.01)	0.22 (-0.32, 0.76)	0.26 (-0.57, 1.09)	0.23 (-0.42, 0.89)	0.25 (-0.32, 0.82)	0.29 (-0.43, 1.01)	0.29 (-0.23, 0.81)	0.32 (-0.30, 0.93)	0.40 (-0.28, 1.07)	0.50 (-0.11, 1.11)	0.63 (-0.33, 1.58)	**0.52 (0.02, 1.02)**	**0.58 (0.02, 1.15)**	**0.66 (0.01, 1.31)**	**0.85 (0.09, 1.61)**	-
9.02 (1.48, 55.09)	1.40 (0.19, 10.19)	2.84 (0.72, 11.25)	OLA	0.04 (-0.76, 0.85)	0.13 (-0.44, 0.71)	0.20 (-0.39, 0.78)	0.17 (-0.69, 1.03)	0.21 (-0.37, 0.80)	0.24 (-0.61, 1.10)	0.22 (-0.47, 0.92)	0.24 (-0.37, 0.85)	0.28 (-0.48, 1.03)	0.28 (-0.28, 0.84)	0.31 (-0.34, 0.96)	0.39 (-0.32, 1.10)	0.49 (-0.15, 1.14)	0.62 (-0.36, 1.59)	0.51 (-0.03, 1.06)	0.57 (-0.03, 1.18)	0.65 (-0.04, 1.33)	**0.84 (0.05, 1.63)**	-
7.83 (1.31, 46.66)	1.22 (0.17, 8.65)	2.47 (0.64, 9.45)	0.87 (0.20, 3.71)	BRO	0.09 (-0.53, 0.71)	0.15 (-0.48, 0.78)	0.13 (-0.77, 1.02)	0.17 (-0.46, 0.80)	0.20 (-0.68, 1.09)	0.18 (-0.55, 0.91)	0.20 (-0.45, 0.85)	0.24 (-0.55, 1.02)	0.24 (-0.38, 0.85)	0.26 (-0.43, 0.96)	0.35 (-0.40, 1.09)	0.45 (-0.24, 1.14)	0.57 (-0.43, 1.58)	0.47 (-0.12, 1.06)	0.58 (-0.12, 1.18)	0.60 (-0.12, 1.33)	0.80 (-0.02, 1.62)	-
7.20 (1.62, 32.07)	1.12 (0.20, 6.13)	2.27 (0.96, 5.36)	0.80 (0.27, 2.34)	0.92 (0.33, 2.58)	PAR	0.06 (-0.20, 0.33)	0.04 (-0.65, 0.73)	0.08 (-0.19, 0.35)	0.11 (-0.57, 0.79)	0.09 (-0.37, 0.55)	0.11 (-0.21, 0.42)	0.15 (-0.40, 0.69)	0.15 (-0.08, 0.37)	0.18 (-0.22, 0.57)	0.26 (-0.23, 0.74)	0.36 (-0.03, 0.75)	0.48 (-0.34, 1.31)	**0.38 (0.21, 0.55)**	**0.44 (0.12, 0.75)**	**0.51 (0.07, 0.96)**	**0.71 (0.11, 1.31)**	-
6.25 (1.39, 28.19)	0.97 (0.18, 5.38)	1.97 (0.77, 5.05)	0.69 (0.23, 2.07)	0.80 (0.28, 2.28)	0.87 (0.59, 1.28)	VEN	-0.02 (-0.72, 0.67)	0.02 (-0.27, 0.30)	0.05 (-0.64, 0.74)	0.03 (-0.44, 0.50)	0.04 (-0.29, 0.37)	0.08 (-0.47, 0.64)	0.08 (-0.14, 0.30)	0.11 (-0.29, 0.52)	0.19 (-0.30, 0.69)	0.30 (-0.10, 0.70)	0.42 (-0.41, 1.26)	**0.32 (0.12, 0.52)**	**0.38 (0.04, 0.71)**	0.45 (-0.01, 0.91)	**0.65 (0.05, 1.25)**	-
11.29 (1.57, 81.20)	1.76 (0.21, 14.83)	3.56 (0.73, 17.35)	1.25 (0.23, 6.70)	1.44 (0.28, 7.52)	1.57 (0.41, 5.93)	1.81 (0.47, 6.93)	AMI	0.04 (-0.66, 0.74)	0.08 (-0.86, 1.01)	0.05 (-0.74, 0.84)	0.07 (-0.65, 0.79)	0.11 (-0.74, 0.95)	0.11 (-0.57, 0.79)	0.14 (-0.62, 0.89)	0.22 (-0.59, 1.02)	0.32 (-0.43, 1.08)	0.45 (-0.60, 1.50)	0.34 (-0.32, 1.01)	0.40 (-0.32, 1.12)	0.48 (-0.31, 1.26)	0.67 (-0.21, 1.55)	-
5.55 (1.18, 26.04)	0.86 (0.15, 4.95)	1.75 (0.64, 4.77)	0.62 (0.20, 1.93)	0.71 (0.23, 2.14)	0.77 (0.46, 1.29)	0.89 (0.51, 1.54)	0.49 (0.12, 1.97)	FLX	0.04 (-0.66, 0.73)	0.01 (-0.46, 0.49)	0.03 (-0.31, 0.36)	0.07 (-0.49, 0.63)	0.07 (-0.18, 0.32)	0.10 (-0.31, 0.51)	0.18 (-0.32, 0.67)	0.28 (-0.12, 0.69)	0.40 (-0.43, 1.24)	**0.30 (0.09, 0.51)**	**0.36 (0.02, 0.69)**	0.43 (-0.03, 0.90)	**0.63 (0.02, 1.24)**	-
4.09 (1.04, 16.15)	0.64 (0.08, 5.27)	1.29 (0.27, 6.12)	0.45 (0.09, 2.37)	0.52 (0.10, 2.65)	0.57 (0.16, 2.08)	0.65 (0.18, 2.43)	0.36 (0.06, 2.26)	0.74 (0.19, 2.86)	IMI	-0.02 (-0.81, 0.76)	-0.01 (-0.72, 0.70)	0.03 (-0.64, 0.70)	0.03 (-0.69, 0.81)	0.06 (-0.66, 0.94)	0.14 (-0.50, 0.99)	0.25 (-0.67, 1.41)	0.37 (-0.39, 0.92)	0.27 (-0.39, 1.04)	0.33 (-0.38, 1.18)	0.40 (-0.28, 1.47)	0.59	-
11.75 (2.15, 64.07)	1.83 (0.28, 11.98)	**3.70 (1.09, 12.57)**	1.30 (0.34, 4.98)	1.50 (0.41, 5.55)	1.63 (0.68, 3.89)	1.88 (0.77, 4.58)	1.04 (0.22, 4.92)	2.12 (0.81, 5.50)	2.87 (0.62, 13.22)	TPM	0.02 (-0.49, 0.52)	0.05 (-0.62, 0.73)	0.05 (-0.39, 0.50)	0.08 (-0.47, 0.64)	0.17 (-0.46, 0.79)	0.27 (-0.28, 0.82)	0.39 (-0.52, 1.31)	0.29 (-0.14, 0.71)	0.35 (-0.15, 0.85)	0.42 (-0.17, 1.02)	0.62 (-0.09, 1.33)	-
12.81 (2.32, 70.73)	1.99 (0.30, 13.21)	4.04 (0.30, 13.94)	1.42 (0.37, 5.52)	1.64 (0.44, 6.15)	1.78 (0.73, 4.35)	2.05 (0.82, 5.11)	1.13 (0.24, 5.44)	2.31 (0.87, 6.13)	3.13 (0.67, 14.61)	1.09 (0.33, 3.63)	RIS	0.04 (-0.54, 0.62)	0.04 (-0.26, 0.34)	0.07 (-0.37, 0.51)	0.15 (-0.37, 0.67)	0.25 (-0.18, 0.69)	0.38 (-0.47, 1.23)	**0.27 (0.01, 0.54)**	0.33 (-0.04, 0.71)	0.41 (-0.09, 0.90)	0.60 (-0.02, 1.23)	-
10.45 (1.48, 73.64)	1.63 (0.20, 13.47)	3.29 (0.69, 15.65)	1.16 (0.22, 6.06)	1.33 (0.26, 6.79)	1.45 (0.40, 5.33)	1.67 (0.45, 6.23)	0.92 (0.15, 5.77)	1.88 (0.48, 7.33)	2.55 (0.42, 15.56)	0.89 (0.19, 4.10)	0.82 (0.17, 3.81)	NEF	0.00 (-0.52, 0.52)	0.03 (-0.60, 0.66)	0.11 (-0.58, 0.80)	0.22 (-0.41, 0.84)	0.34 (-0.63, 1.30)	0.23 (-0.29, 0.76)	0.29 (-0.29, 0.88)	0.37 (-0.30, 1.04)	0.56 (-0.20, 1.33)	-
8.32 (1.87, 37.08)	1.29 (0.24, 7.09)	**2.62 (1.05, 6.59)**	0.92 (0.32, 2.70)	1.06 (0.38, 2.99)	1.16 (0.83, 1.61)	1.33 (0.95, 1.87)	0.74 (0.19, 2.79)	1.50 (0.90, 2.51)	2.03 (0.56, 7.45)	0.71 (0.30, 1.69)	0.65 (0.27, 1.59)	0.80 (0.22, 2.93)	SER	0.03 (-0.35, 0.41)	0.11 (-0.36, 0.58)	0.21 (-0.16, 0.59)	0.34 (-0.48, 1.16)	**0.23 (0.09, 0.38)**	0.29 (-0.01, 0.59)	0.37 (-0.07, 0.81)	0.56 (0.00, 1.13)	-
9.04 (1.78, 45.88)	1.41 (0.23, 8.64)	2.85 (0.93, 8.74)	1.00 (0.29, 3.50)	1.15 (0.34, 3.89)	1.26 (0.61, 2.58)	1.45 (0.69, 3.05)	0.80 (0.18, 3.50)	1.63 (0.72, 3.70)	2.21 (0.52, 9.39)	0.77 (0.26, 2.26)	0.71 (0.24, 2.12)	0.86 (0.20, 3.68)	1.09 (0.53, 2.23)	NK1R	0.08 (-0.49, 0.66)	0.19 (-0.31, 0.68)	0.31 (-0.58, 1.19)	0.20 (-0.15, 0.56)	0.26 (-0.18, 0.71)	0.34 (-0.21, 0.89)	0.54 (-0.14, 1.21)	-
12.91 (2.10, 79.49)	2.01 (0.27, 14.70)	**4.07 (1.02, 16.27)**	1.43 (0.32, 6.37)	1.65 (0.38, 7.12)	1.79 (0.61, 5.32)	2.07 (0.69, 6.23)	1.14 (0.21, 6.17)	2.33 (0.73, 7.39)	3.16 (0.60, 16.60)	1.10 (0.28, 4.25)	1.01 (0.26, 3.96)	1.24 (0.23, 6.52)	1.55 (0.52, 4.60)	1.43 (0.41, 5.04)	GUA	0.10 (-0.47, 0.68)	0.23 (-0.70, 1.16)	0.12 (-0.33, 0.58)	0.18 (-0.34, 0.71)	0.26 (-0.36, 0.87)	0.45 (-0.27, 1.18)	-
4.68 (0.97, 22.43)	0.73 (0.12, 4.25)	1.47 (0.52, 4.16)	0.52 (0.16, 1.68)	0.60 (0.19, 1.86)	0.65 (0.36, 1.16)	0.75 (0.41, 1.38)	0.41 (0.10, 1.70)	0.84 (0.42, 1.70)	1.14 (0.29, 4.56)	0.40 (0.15, 1.07)	0.37 (0.13, 1.00)	0.45 (0.11, 1.79)	0.56 (0.31, 1.00)	0.52 (0.22, 1.23)	0.36 (0.11, 1.19)	TGB	0.12 (-0.76, 1.00)	0.02 (-0.33, 0.37)	0.08 (-0.51, 0.36)	0.15 (-0.39, 0.70)	0.35 (-0.32, 1.02)	-
10.00 (0.94, 106.64)	1.56 (0.13, 18.99)	3.15 (0.40, 24.58)	1.11 (0.13, 9.31)	1.28 (0.16, 10.50)	1.39 (0.22, 8.97)	1.60 (0.25, 10.44)	0.89 (0.09, 8.55)	1.80 (0.27, 12.13)	2.44 (0.26, 23.15)	0.85 (0.11, 6.49)	0.78 (0.10, 6.01)	0.96 (0.10, 9.08)	1.20 (0.19, 7.76)	1.11 (0.15, 7.95)	0.77 (0.09, 6.54)	2.14 (0.31, 14.67)	BUP	-0.10 (-0.91, 0.71)	0.05 (-0.81, 0.90)	0.03 (-0.88, 0.94)	0.23 (-0.76, 1.22)	-
7.50 (1.72, 32.80)	1.17 (0.22, 6.28)	2.36 (0.97, 5.76)	0.83 (0.29, 2.37)	0.96 (0.35, 2.62)	1.04 (0.62, 1.32)	1.20 (0.89, 1.63)	0.66 (0.18, 2.46)	1.35 (0.85, 2.14)	1.83 (0.51, 6.57)	0.64 (0.28, 1.48)	0.59 (0.25, 1.39)	0.72 (0.20, 2.58)	0.90 (0.71, 1.14)	0.82 (0.42, 1.64)	0.75 (0.20, 1.68)	1.68 (0.94, 2.73)	0.75 (0.12, 4.77)	PLB	0.06 (-0.20, 0.32)	0.13 (-0.28, 0.55)	0.33 (-0.24, 0.90)	-
7.87 (1.66, 37.31)	1.22 (0.21, 7.08)	2.48 (0.89, 6.87)	0.87 (0.27, 2.78)	1.01 (0.33, 3.08)	1.09 (0.63, 1.89)	1.26 (0.70, 2.25)	0.70 (0.17, 2.83)	1.42 (0.72, 2.79)	1.92 (0.49, 7.56)	0.67 (0.25, 1.77)	0.61 (0.23, 1.66)	0.75 (0.19, 2.97)	0.94 (0.55, 1.64)	0.87 (0.37, 2.02)	0.61 (0.19, 1.97)	1.68 (0.81, 3.48)	1.05 (0.64, 1.72)	0.82?	PRZ	0.07 (-0.41, 0.57)	0.27 (-0.36, 0.90)	-
8.21 (1.48, 45.51)	1.28 (0.19, 8.50)	2.59 (0.75, 8.98)	0.91 (0.23, 3.56)	1.05 (0.28, 3.96)	1.14 (0.46, 2.81)	1.31 (0.52, 3.30)	0.73 (0.15, 3.50)	1.48 (0.55, 3.95)	2.01 (0.43, 9.40)	0.70 (0.21, 2.34)	0.64 (0.19, 2.18)	0.79 (0.17, 3.69)	0.99 (0.40, 2.43)	0.91 (0.30, 2.74)	0.64 (0.16, 2.51)	1.76 (0.63, 4.86)	0.82 (0.11, 6.35)	1.10 (0.46, 2.61)	1.04 (0.38, 2.84)	DVP	0.20 (-0.51, 0.90)	-
5.58 (0.81, 38.61)	0.87 (0.11, 7.07)	1.76 (0.38, 8.17)	0.62 (0.12, 3.17)	0.71 (0.14, 3.55)	0.77 (0.22, 2.77)	0.89 (0.25, 3.21)	0.49 (0.08, 3.02)	1.00 (0.26, 3.81)	1.36 (0.23, 8.15)	0.47 (0.11, 2.14)	0.44 (0.10, 1.99)	0.53 (0.09, 3.20)	0.67 (0.19, 2.32)	0.62 (0.15, 2.57)	0.43 (0.08, 2.23)	1.19 (0.31, 4.65)	0.56 (0.06, 5.21)	0.74 (0.21, 2.60)	0.71 (0.18, 2.73)	0.68 (0.15, 3.12)	CIT	0.17 (0.01, 3.08)	-
0.94 (0.05, 19.02)	0.15 (0.01, 3.29)	0.30 (0.02, 4.72)	0.10 (0.01, 1.75)	0.12 (0.01, 1.99)	0.13 (0.01, 1.81)	0.15 (0.01, 2.10)	0.08 (0.00, 1.56)	0.17 (0.01, 2.42)	0.23 (0.01, 4.24)	0.08 (0.01, 1.25)	0.07 (0.00, 1.16)	0.09 (0.00, 1.66)	0.11 (0.01, 1.57)	0.10 (0.01, 1.56)	0.07 (0.00, 1.23)	0.20 (0.01, 2.92)	0.09 (0.00, 2.32)	0.12 (0.01, 1.72)	0.12 (0.01, 1.72)	0.11 (0.01, 1.81)	0.17 (0.01, 3.08)	LAM	-

Legend: Mean change in symptoms severity (green) and all-cause dropout rate (blue). For efficacy (green), standardized mean difference values greater than 0 indicated that the treatment specified in the row is more efficacious. For acceptability, odds ratios greater than 1 indicated that the treatment specified in the column is better (fewer dropouts). Bold underlined results indicate statistical significance. The overall heterogeneity (τ) is equal to 0.1 for efficacy and equal to 0 for acceptability. For efficacy, I^2 was 22.4% for heterogeneity and 22.5% for inconsistency; for acceptability, I^2 was 0% for heterogeneity and 8% for inconsistency. AMI, amitriptyline; BRO, bromipramine; BUP, bupropion; CIT, citalopram; DES, desipramine; DVP, divalproex; FLX, fluoxetine; GUA, guanfacine; IMI, imipramine; LAM, lamotrigine; MIR, mirtazapine; NEF, nefazodone; NK1R, NK1 receptor antagonist; OLA, olanzapine; PAR, paroxetine; PHE, phenelzine; PLB, placebo; PRZ, prazosin; RIS, risperidone; SER, sertraline; TGB, tiagabine; TPM, topiramate; VEN, venlafaxine.

In Fig. 4 we used the hierarchical cluster analysis to group treatments according to their ranking for the two primary outcomes. Different colours represent different groups of treatments by considering jointly their relative ranking for two outcomes (treatments that belong to the same group may be considered as being of comparable performance with respect to both outcomes). According to all possible cluster rankings, phenelzine performed better than the rest of the treatments both in terms of efficacy and dropouts (Appendix 11); this was also supported by considering the actual effect sizes presented in Table 1 and Fig. 3. Although mirtazapine yielded a relatively high rank in the efficacy outcomes, the respective SUCRA value for the dropout outcome was 51.5% (Appendix 12). Brofaromine performed well in terms of change in symptoms and adverse events, whereas divalproex had the worst ranking considering this particular pair of outcomes. Risperidone was the best treatment in terms of response rate whereas it is among the worst for the outcome of adverse events. See Appendix 12 for all details about rankings and SUCRA values. We could not perform the pre-planned sensitivity analyses because only five studies were rated as low risk of bias for allocation concealment (Appendix 3) and because we did not impute any outcome data (all information we needed was retrieved by checking for unpublished data and contacting the original study authors). According to GRADE, the quality of evidence for primary outcomes was rated as low or very low for most comparisons.

Impact of effect modifiers

Because of concerns about power to detect potentially important heterogeneity and inconsistency, we applied network meta-regression models to explore the impact on effect modifiers on heterogeneity and inconsistency. We included the pre-specified covariates in separate network meta-regression models. All results are presented in Appendix 13. For the primary efficacy outcome, the meta-regression model that accounted for baseline severity resulted in a marginally non-significant coefficient 0.11 (95% CI −0.01 to 0.22) implying that the active treatments tended to be more effective (compared with placebo) in studies with more severe patients. The estimation of heterogeneity (τ) was of 0.1 for both the meta-regression model accounting for baseline severity and the model without covariates, showing that accounting for baseline severity did not explain heterogeneity. Appendix 14 presents the effect sizes and the treatment hierarchy (using SUCRAs) of the standard analysis and those from the meta-regression model at the mean centralised baseline severity. Accounting for differences in baseline severity across studies had a minimal impact on the effect sizes and the treatment ranking.

The funnel plot analyses suggested a possible association between study precision and dropout due to adverse events (Appendix 15). However, the relative effect estimates derived from the network meta-regression model that accounted for

(a) Change in symptoms

Drug	SMD (95%CI)
Phenelzine	-0.97 (-1.68, -0.27)
Mirtazapine	-0.79 (-1.66, 0.08)
Desipramine	-0.52 (-1.03, -0.02)
Olanzapine	-0.51 (-1.05, 0.03)
Brofaromine	-0.47 (-1.07, 0.13)
Paroxetine	-0.38 (-0.55, -0.20)
Amitriptyline	-0.34 (-1.01, 0.32)
Venlafaxine	-0.32 (-0.53, -0.11)
Fluoxetine	-0.30 (-0.51, -0.09)
Topiramate	-0.29 (-0.71, 0.14)
Risperidone	-0.28 (-0.54, -0.01)
Imipramine	-0.27 (-0.92, 0.39)
Sertraline	-0.24 (-0.38, -0.09)
Nefazodone	-0.24 (-0.76, 0.29)
NK1R antagonist	-0.21 (-0.56, 0.15)
Guanfacine	-0.12 (-0.58, 0.33)
Tiagabine	-0.02 (-0.38, 0.34)
Prazosin	0.06 (-0.21, 0.33)
Bupropion	0.10 (-0.70, 0.91)
Divalproex	0.14 (-0.28, 0.55)
Citalopram	0.33 (-0.24, 0.90)

-1.5 -1 -.5 0 .5 1
Favors treatment   Favors placebo

(b) All-cause dropouts

Drug	OR (95%CI)
Lamotrigine	0.12 (0.01, 1.72)
Phenelzine	0.13 (0.03, 0.58)
Desipramine	0.42 (0.17, 1.03)
Imipramine	0.55 (0.15, 1.96)
Tiagabine	0.62 (0.37, 1.06)
Fluoxetine	0.74 (0.47, 1.17)
Citalopram	0.74 (0.21, 2.60)
Venlafaxine	0.83 (0.61, 1.13)
Mirtazapine	0.86 (0.16, 4.62)
Paroxetine	0.96 (0.76, 1.21)
Brofaromine	1.04 (0.38, 2.85)
Prazosin	1.05 (0.64, 1.72)
Divalproex	1.10 (0.46, 2.61)
Sertraline	1.11 (0.88, 1.40)
Olanzapine	1.20 (0.42, 3.43)
NK1R antagonist	1.20 (0.61, 2.38)
Buproprion	1.33 (0.21, 8.49)
Nefazodone	1.39 (0.39, 5.01)
Amitriptyline	1.51 (0.41, 5.58)
Topiramate	1.57 (0.68, 3.62)
Risperidone	1.71 (0.72, 4.04)
Guanfacine	1.72 (0.60, 4.97)

0.1   0.5 1 2   5 10
Favors treatment   Favors placebo

Fig. 3. Network meta-analysis results for the comparisons of active treatments *v.* placebo for the primary outcomes (a) change in symptoms and (b) all-cause dropouts (reference treatment: placebo). Legend: SMD: standardised mean difference; OR: odds ratio; CI: confidence interval. Effect sizes are listed on the basis of SUCRA rankings. Statistically significant results are highlighted in red.
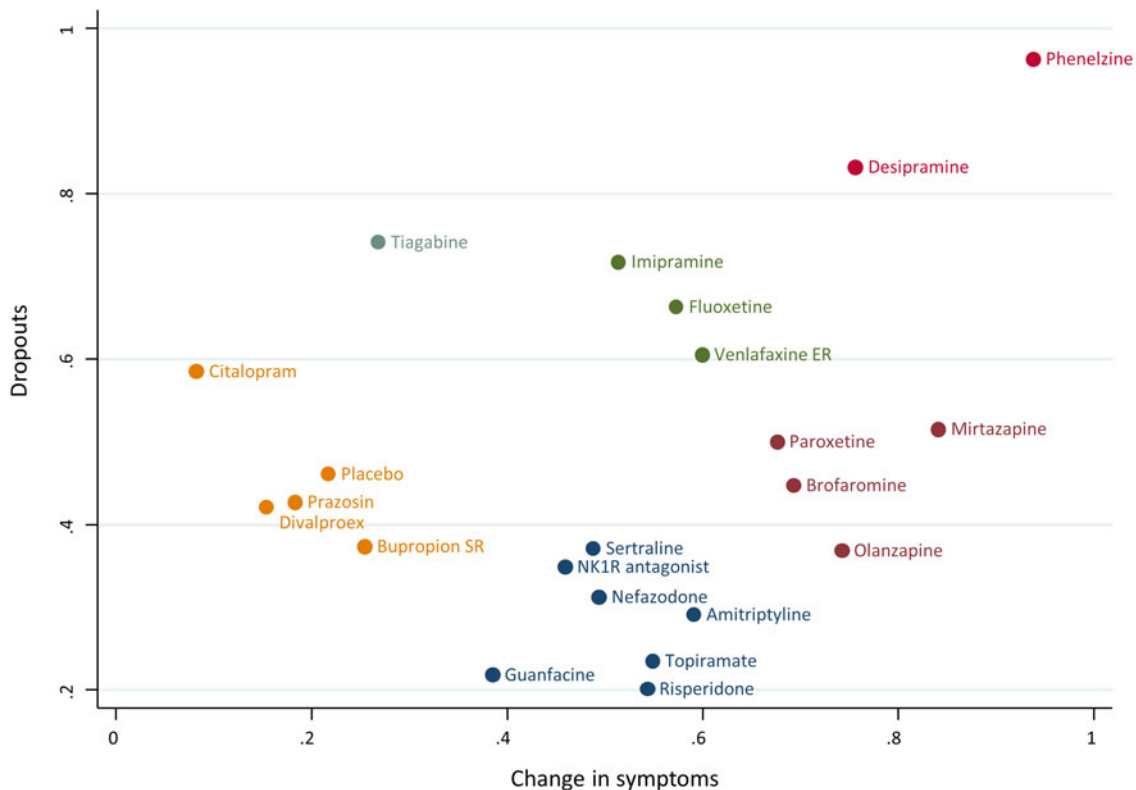
Fig. 4. Clustered ranking based on SUCRA values for the outcomes 'Change in symptoms' (unadjusted model) and 'Any cause dropouts'. Hierarchical cluster analysis is performed to group the competing treatments. Different colours represent different groups according to treatments' similarity regarding the two outcomes.

small study effects did not converge and drawing conclusions based on these findings would be invalid. We performed a random-effects pairwise meta-analysis for the comparison active v. placebo which resulted in a summary effect OR of 1.42 (95% CI 1.15–1.75) indicating that placebo is associated with fewer dropouts due to adverse events than active treatments. Accounting for study precision in a meta-regression model resulted in a statistically significant coefficient 2.07 (95% CI 1.10–3.88). This result suggests that less precise studies tended to show placebo to be better tolerated than it would have been in larger studies (Appendix 15).

Discussion

Our network meta-analysis provides an evidence-based hierarchy for the efficacy and acceptability of pharmacological treatments for PTSD, using all available comparative data and thus overcoming the major limitation of previous pairwise meta-analyses (NICE, 2005; Watts et al. 2013; Hoskins et al. 2015). Interestingly, results from this study challenge the previous dogma that the efficacy of psychopharmacological agents within the same drug class is the same. Similarly, clinically important differences in efficacy were found among antipsychotics for schizophrenia (Leucht et al. 2013) and also among antidepressants, where differences were reported not only in major depression (Cipriani et al. 2009), but also for anxiety disorders, like generalised anxiety disorder (NICE, 2011) and social anxiety disorder (NICE, 2013). The efficacy and acceptability hierarchies generated by our study were robust against many sources of bias, as various supplementary analyses related to severity at baseline, year of publication, small study effect, and sponsorship did not change the final ranking to an important extent.

We emphasise that the statistically significant differences in efficacy between drugs and placebo were small, with the only exception of phenelzine (according to Cohen, 1988, an effect size of 0.2 is small, 0.5 is medium, and 0.8 is large (Cohen, 1988)). It is probably misleading clinically to dismiss the effects of desipramine, fluoxetine, paroxetine, risperidone, sertraline, and venlafaxine as being small. Their point estimates of SMD were around 0.3, which is the same as antidepressants for depression, and indeed for many standard treatments in psychiatry and medicine (Leucht et al. 2012). Phenelzine had a standardised mean difference of 0.97 when compared with placebo and ranged between 0.19 and 1.30 with other active drugs. These differences in efficacy between active treatments were possibly substantial enough to be clinically important and pose the question whether phenelzine should be used as a first line drug treatment for PTSD. However, the overall evidence about this drug is based only on one trial with 60 participants. We assessed the contribution of this single study to all 'phenelzine versus other' network estimates for the primary outcome 'change in symptoms' (Appendix 16 - Fig. 1). The only study comparing phenelzine contributes more than 35% to the relative effects of all treatments against this specific intervention. Given that the contribution of this study to the relative effects of the rest of the treatment comparisons is very low (results not shown but available from the authors), it is therefore unlikely that the retrieved evidence can currently be considered robust enough to suggest phenelzine as a drug of choice, particularly in view of the need for concomitant dietary restrictions and the risk of drug interactions. This is in line with the interpretation of recent clinical practice guidelines (Katzman et al. 2014),

however, our analysis had higher statistical power than previous analyses, because of the comprehensive search of published and unpublished data and because we used the network meta-analytical method.

Phenelzine is a potent inhibitor of monoamine oxidase (MAOI), licensed in the USA and UK for depressed patients clinically characterised as 'atypical,' 'non-endogenous,' or 'neurotic', who often have mixed anxiety and depression and phobic or hypochondriacal features (http://www.accessdata.fda.gov/drugsatfda_docs/label/2007/011909s038lbl.pdf). Even though approved for depression only, phenelzine is also effective in the treatment of anxiety disorders such as panic disorder and social phobia (Sheehan et al. 1980; Johnson et al. 1994). Of course, phenelzine may owe its apparently striking efficacy in PTSD to the expected increases in brain levels of biogenic amines (serotonin, noradrenaline, and dopamine) produced by MAO inhibition. However, the reversible MAOI-A inhibitor, brofaromine, was of only modest efficacy in the current analysis. It may, therefore, be relevant that unlike other MAOIs, phenelzine also possesses the ability (at least in animal studies) to increase brain levels of the inhibitory neurotransmitter, γ-aminobutyric acid (GABA), through blocking the GABA metabolising enzyme, GABA transaminase. Studies with magnetic resonance spectroscopy (MRS) indicate that cortical GABA levels may be lower in patients with PTSD (Rosso et al. 2014). Hence the efficacy of phenelzine in PTSD might result from a combination of monoamine and GABA potentiation. However, it is also possible that GABA transaminase inhibition by itself, for example with the anticonvulsant drug vigabatrin, could be a viable therapeutic strategy in PTSD.

Interestingly, treatment with phenelzine reduces dream recall frequency (Landolt et al. 2001) and induces almost total REM sleep suppression in anxious-depressed patients (Wyatt et al. 1971). Although objective sleep findings in PTSD are mixed, there is some evidence that REM sleep fragmentation and REM sleep autonomic imbalance are associated with an increased risk for the development and persistence of PTSD (Mellman et al. 2007). Sleep disturbances are an indicator of heightened risk for poor psychiatric outcomes following trauma exposure (Germain, 2013). To date, prazosin (an alpha-1 antagonist) is the recommended pharmacological treatment options for PTSD-related nightmares, as it is associated with improvements in nightmares and insomnia (Aurora et al. 2010). In our analysis prazosin was not among the most efficacious treatments for relieving acute symptoms of PTSD, so the fact that phenelzine may reduce the REM sleep disturbances associated with PTSD could also give it additional benefit over drugs like SSRIs and SNRIs which also lower REM sleep, but to a lesser extent than phenelzine (Sharpley & Cowen, 1995).

Our study has several limitations. The validity of the above findings might be limited due to the lack of sufficient data to properly evaluate the required assumptions. Results from network meta-analysis may be unreliable when there is a high risk of intransitivity in the network (Salanti, 2012; Cipriani et al. 2013). Statistical assessment of the transitivity assumption (i.e. one can learn indirectly for a pairwise comparison via one or more anchor treatments) considers whether the distribution of the potential effect modifiers is balanced across the available direct comparisons. In our network, we were unable to judge the plausibility of transitivity since most comparisons included very few studies. Thus, we could only rely on our clinical understanding of the studied condition and outcomes to assume that transitivity was likely to hold in the identified data. Important intransitivity might be

reflected in the data in the form of statistical inconsistency. Our network included only a few closed loops of evidence, for which inconsistency could be assessed (three loops for efficacy as a continuous outcome, two for all-cause dropout, none for response rate and two for dropouts due to adverse events). Using the loop-specific approach and the design-by-treatment interaction model we did not find any statistically significant inconsistency (Song et al. 2003; Higgins et al. 2012). However, the tests for inconsistency very often have low power and may fail to detect inconsistency as statistically significant even when it is present (Veroniki et al. 2013). Inconsistency is closely related with heterogeneity and the two notions should be considered jointly. The small number of studies in the majority of the available direct comparisons renders the assessment of comparison-specific heterogeneities impractical (when it is estimable). The common heterogeneity across all comparisons from the network meta-analysis was estimated close to zero and the I^2 measure for heterogeneity suggested that heterogeneity was low in the network. This finding slightly mitigates our concerns for the reliability of our results. However, all statements comparing the drugs in the network must be tempered by the potential limitations of the methodology, the complexity of patients with PTSD and the uncertainties that may result from the choice of dose or treatment setting.

We explored the association of several characteristics and their impact on the relative effects using network meta-regression models. The model accounting for baseline severity for the efficacy primary outcome resulted in a marginally non-statistically significant regression coefficient suggesting that the active treatments have a tendency to appear more effective (when compared with placebo) in studies with more severe populations. The meta-regression model that accounted for small-study effects on the outcome of adverse events suggested that small studies showed the placebo safer than did larger studies. This finding might be due to publication bias or selective reporting bias, or might be the manifestation of other characteristics such as study quality (Chaimani & Salanti, 2012; Chaimani et al. 2013b). Despite this finding, we were unable to use the results from this model for drawing conclusions with respect to the relative safety of the treatments due to the lack of convergence for the estimated relative effect.

Most trials included in our analysis did not report adequate information about randomisation and allocation concealment, and this might undermine the validity of overall findings. However, all the studies were double-blind and the scant information in terms of quality assessment could be more an issue of reporting in the text than real defects in study design, as it has been commonly found in other systematic reviews (Huwiler-Müntener et al. 2002). Finally, findings from this analysis apply only to acute-phase treatment of PTSD. Clinically, the assessment of efficacy after 4 weeks of treatment or after 16 weeks or more might lead to wide differences in treatment outcome, and this may limit the ability to provide valid estimates of treatment effect if different durations of follow-up are combined. To overcome this problem, we employed a common definition of acute response that included a pre-defined follow-up duration (8 weeks). Moreover, because network meta-analysis requires reasonably homogeneous studies, in this project we had to restrict ourselves to short-term trials and decided to address the issue of long-term relapse prevention treatment in an ad hoc network meta-analysis, as we previously did for bipolar disorder (Cipriani et al. 2011; Miura et al. 2014).

All-cause discontinuation has been used in other NMAs as a measure of the acceptability of treatments because it encompasses at the same time efficacy and tolerability. Response rate and dropout rate were chosen as primary outcomes to have the best estimate of acute treatment effectiveness (Cipriani et al. 2009; Leucht et al. 2013). In our analysis, the results paralleled the efficacy findings in that phenelzine, the most effective drug also had the lowest discontinuation rates. We used the neutral term all-cause discontinuation because clinicians might intuitively associate the word acceptability more with tolerability than with efficacy. However, it is worth noting that phenelzine, even though not statistically significant, showed a trend of better tolerability in terms of dropout rate due to adverse events when compared with placebo. Even if in clinical practice MAOIs are used much less frequently than other antidepressant agents because of the dangers of dietary and drug interactions (http://www.bnf.org/bnf/index.htm), findings from this review reinforce the idea that phenelzine should be prioritised in future trials in PTSD, particularly in the significant proportion of patients who remain symptomatic despite first-line psychological and pharmacological treatments. It would also be of interest to use magnetic resonance spectroscopy to assess the effect of phenelzine treatment on GABA levels in the human brain and to relate symptomatic improvement in PTSD symptomatology to changes in GABA concentration.

References

Altman DG and Bland JM (1996) Detecting skewness from summary information. *British Medical Journal* **313**, 1200.

American Psychiatric Association (2004) *Practice Guideline for the Treatment of Patients with Acute Stress Disorder and Posttraumatic Stress Disorder*. Arlington: American Psychiatric Publishing.

Aurora RN, Zak RS, Auerbach SH, Casey KR, Chowdhuri S, Karippot A et al. Standards of Practice Committee; American Academy of Sleep Medicine (2010) Best practice guide for the treatment of nightmare disorder in adults. *Journal of Clinical Sleep Medicine* **6**, 389–401.

Chaimani A, Higgins JPT, Mavridis D, Spyridonos P and Salanti G (2013a) Graphical tools for network meta-analysis in STATA. *PLoS ONE* **8**, e76654.

Chaimani A and Salanti G (2012) Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. *Research Synthesis and Methods* **3**, 161–176.

Chaimani A, Vasiliadis HS, Pandis N, Schmid CH, Welton NJ and Salanti G (2013*b*) Effects of study precision and risk of bias in networks of interventions: a network meta-epidemiological study. *International Journal of Epidemiology* 42, 1120–1131.

Cipriani A, Barbui C, Salanti G, Rendell J, Brown R, Stockton S *et al.* (2011) Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments meta-analysis. *Lancet* 378, 1306–1315.

Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R *et al.* (2009) Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 373, 746–758.

Cipriani A, Higgins JP, Geddes JR and Salanti G (2013) Conceptual and technical challenges in network meta-analysis. *Annals of Internal Medicine* 159, 130–137.

Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Hillsdale: Lawrence Erlbaum Associates.

DerSimonian R and Laird N (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials* 7, 177–188.

Dohrenwend BP, Turner JB, Turse NA, Adams BG, Koenen KC and Marshall R (2006) The psychological risks of Vietnam for U.S. veterans: a revisit with new data and methods. *Science* 313, 979–982.

Furukawa TA, Barbui C, Cipriani A, Brambilla P and Watanabe N (2006) Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology* 59, 7–10.

Germain A (2013) Sleep disturbances as the hallmark of PTSD: where are we now? *American Journal of Psychiatry* 170, 372–382.

Higgins JP and Green S (2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated September 2011]. Available at http://www.cochrane-handbook.org.

Higgins JP, Thompson SG, Deeks JJ and Altman DG (2003) Measuring inconsistency in meta-analyses. *British Medical Journal* 327, 557–560.

Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE and White IR (2012) Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Research Synthesis and Methods* 3, 98–110.

Hoskins M, Pearce J, Bethell A, Dankova L, Barbui C, Tol WA *et al.* (2015) Pharmacotherapy for post-traumatic stress disorder: systematic review and meta-analysis. *British Journal of Psychiatry* 206, 93–100.

Huwiler-Müntener K, Jüni P, Junker C and Egger M (2002) Quality of reporting of randomized trials as a measure of methodologic quality. *Journal of American Medical Association* 287, 2801–2804.

Jackson D, Barrett JK, Stephen R, White IR and Higgins JPT (2014) A design-by-treatment interaction model for network meta-analysis with random inconsistency effects. *Statistics in Medicine* 33, 3639–3654.

Jansen JP and Naci H (2013) Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Medicine* 11, 159.

Johnson MR, Lydiard RB and Ballenger JC (1994) MAOIs in panic disorder and agoraphobia. In Kennedy SH (ed.). *Clinical Advances in Monoamine Oxidase Inhibitor Therapies*. Washington, DC: American Psychiatric Press, pp. 205–224.

Jonas DE, Cusack K, Forneris CA, Wilkins TM, Sonis J, Middleton JC *et al.* (2013) Psychological and Pharmacological Treatments for Adults With Posttraumatic Stress Disorder (PTSD). Rockville (MD): Agency for Healthcare Research and Quality (US). Report No.: 13-EHC011-EF.

Katzman MA, Bleau P, Blier P, Chokka P, Kjernisted K, Van Ameringen M; Canadian Anxiety Guidelines Initiative Group on behalf of the Anxiety Disorders Association of Canada/Association Canadienne des troubles anxieux and McGill University *et al.* (2014) Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. BMC Psychiatry 14, S1.

Kessler RC, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Anthony JC and Brugha TS *et al.* (2011) Prevalence and severity of mental disorders in the world mental health survey initiative. In Kessler RC and Ustun B (eds). *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*. Cambridge: Cambridge University Press, pp. 511–521.

Krysinska K and Lester D (2010) Post-traumatic stress disorder and suicide risk: a systematic review. *Archives of Suicide Research* 14, 1–23.

Landolt HP, Raimo EB, Schnierow BJ, Kelsoe JR, Rapaport MH and Gillin JC (2001) Sleep and sleep electroencephalogram in depressed patients treated with phenelzine. *Archives of General Psychiatry* 58, 268–276.

Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F *et al.* (2013) Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* 382, 951–962.

Leucht S, Hierl S, Kissling W, Dold M and Davis JM (2012) Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. *British Journal of Psychiatry* 200, 97–106.

Mellman TA, Pigeon WR, Nowell PD and Nolan B (2007) Relationships between REM sleep findings and PTSD symptoms during the early aftermath of trauma. *Journal of Traumatic Stress* 20, 893–901.

Miura T, Noma H, Furukawa TA, Mitsuyasu H, Tanaka S, Stockton S *et al.* (2014) Comparative efficacy and acceptability of pharmacological treatments in the maintenance treatment of bipolar disorder: a network meta-analysis. *Lancet Psychiatry* 1, 351–359.

National Institute for Health and Care Excellence (2005) Post-traumatic stress disorder (PTSD). The management of PTSD in adults and children in primary and secondary care. NICE clinical guideline 26. Available at http://www.nice.org.uk/guidance/cg26/resources/guidance-posttraumatic-stress-disorder-ptsd-pdf.

National Institute for Health and Care Excellence (2011) Generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults. Management in primary, secondary and community care. NICE clinical guideline 113. Available at http://www.nice.org.uk/guidance/cg113/resources/guidance-generalised-anxiety-disorder-and-panic-disorder-with-or-without-agoraphobia-in-adults-pdf.

National Institute for Health and Care Excellence (2013) Social anxiety disorder: recognition, assessment and treatment. NICE clinical guideline 159. Available at http://www.nice.org.uk/guidance/cg159/resources/guidance-social-anxiety-disorder-recognition-assessment-and-treatment-pdf.

Norris F and Sloane LB (2007) The epidemiology of trauma and PTSD. In Friedman MJ, Keane TM and Resick PA (eds). *Handbook of PTSD: Science and Practice*. New York: Guilford Press, pp. 78–98.

Phoenix Australia - Centre for Posttraumatic Mental Health (2013) *Australian Guidelines for the Treatment of Acute Stress Disorder and Posttraumatic Stress Disorder*. Melbourne, Victoria: Phoenix Australia. Available at http://phoenixaustralia.org/wp-content/uploads/2015/03/Phoenix-ASD-PTSD-Guidelines.pdf.

Rhodes KM, Turner RM and Higgins JP (2015) Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *Journal of Clinical Epidemiology* 68, 52–60.

Rosso IM, Weiner MR, Crowley DJ, Silveri MM, Rauch SL and Jensen JE (2014) Insula and anterior cingulate GABA levels in posttraumatic stress disorder: preliminary findings using magnetic resonance spectroscopy. *Depression and Anxiety* 31, 115–123.

Salanti G (2012) Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis and Methods* 3, 80–97.

Salanti G, Ades AE and Ioannidis JP (2011) Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *Journal of Clinical Epidemiology* 64, 163–171.

Salanti G, Del Giovane C, Chaimani A, Caldwell DM and Higgins JP (2014) Evaluating the quality of evidence from a network meta-analysis. *PLoS ONE* 9, e99682.

Seal KH, Bertenthal D, Miner CR, Sen S and Marmar C (2007) Bringing the war back home: mental health disorders among 103788 US veterans returning from Iraq and Afghanistan seen at Department of Veterans Affairs facilities. *Archives of Internal Medicine* 167, 476–482.

Sharpley AL and Cowen PJ (1995) Effect of pharmacologic treatments on the sleep of depressed patients. *Biological Psychiatry* 37, 85–98.

Sheehan DV, Ballenger J and Jacobsen G (1980) Treatment of endogenous anxiety with phobic, hysterical and hypochondrial symptoms. *Archives of General Psychiatry* 37, 51–59.

Song F, Altman DG, Glenny AM and Deeks JJ (2003) Validity of indirect comparison for estimating efficacy of competing interventions:

empirical evidence from published meta-analyses. *British Medical Journal* **326**, 472.

StataCorp (2015) *Stata Statistical Software: Release 14.* College Station, TX: StataCorp LP.

Steel Z, Chey T, Silove D, Marnane C, Bryant RA and van Ommeren M (2009) Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement: a systematic review and meta-analysis. *Journal of the American Medical Association* **302**, 537–549.

Turner RM, Davey J, Clarke MJ, Thompson SG and Higgins JP (2012) Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *International Journal of Epidemiology* **41**, 818–827.

Veroniki AA, Vasiliadis HS, Higgins JP and Salanti G (2013) Evaluation of inconsistency in networks of interventions. *International Journal of Epidemiology* **42**, 332–345.

Watts BV, Schnurr PP, Mayo L, Young-Xu Y, Weeks WB and Friedman MJ (2013) Meta-analysis of the efficacy of treatments for posttraumatic stress disorder. *Journal of Clinical Psychiatry* **74**, e541–e550.

White IR (2011) Multivariate random-effects meta-regression: updates to mvmeta. *The STATA Journal* **11**, 255–270.

White IR, Barrett JK, Jackson D and Higgins JPT (2012) Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Research Synthesis and Methods* **3**, 111–125.

Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG *et al.* (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *British Medical Journal* **336**, 601–605.

Wyatt RJ, Fram DH, Kupfer DJ and Snyder F (1971) Total prolonged drug-induced REM sleep suppression in anxious-depressed patients. *Archives of General Psychiatry* **24**, 145–155.