



Comparative efficacy and acceptability of first-line drugs for the acute treatment of generalized anxiety disorder in adults: A network meta-analysis

Hairong He^{a,b}, Yutao Xiang^c, Fengjie Gao^d, Ling Bai^a, Fan Gao^a, Yajuan Fan^d, Jun Lyu^{a,**}, Xiancang Ma^{d,e,f,*}

^a Clinical Research Center, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, 710061, China

^b School of Public Health, Xi'an Jiaotong University Health Science Center, Xi'an, Shaanxi, China

^c Unit of Psychiatry, Faculty of Health Sciences, University of Macau, Macao, SAR, China

^d Department of Psychiatry, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, 710061, China

^e Clinical Research Center for Mental Disease of Shaanxi Province, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, 710061, China

^f Center for Brain Science, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, 710061, China

ARTICLE INFO

Keywords:

Network meta-analysis
Generalized anxiety disorder
First-line treatment drugs

ABSTRACT

The guide recommends SSRI and SNRI drugs as first-line treatments for generalized anxiety disorder (GAD). Therefore, we aimed to update the evidence using network meta-analysis by comparing the efficacy and acceptability of first-line drugs. The relevant electronic databases were searched for placebo-controlled and head-to-head trials of 11 drugs used for the acute treatment of adults with GAD from 1980 up to January 1, 2019. Data on demographics, clinical, and treatment information were extracted from each eligible study. The primary outcomes were efficacy (quantified as the change in the total score on the Hamilton Anxiety Scale from baseline) and acceptability (quantified as treatment discontinuations due to any cause). Overall, the data on 41 RCTs were sufficient or appropriate for inclusion. In terms of efficacy, all of the drugs except fluoxetine and vortioxetine were more effective than placebo, with the weighted mean difference of the Hamilton Anxiety Scale score ranging between -3.2 (95% credible interval [CrI] = -4.2 to -2.2) for escitalopram and -1.8 (95% CrI = -3.1 to -0.55) for vilazodone. For acceptability, only vilazodone (OR = 1.7, 95% CrI = 1.1 to 2.7) were worse than placebo, others did not show significant differences from placebo. In head-to-head comparisons, vortioxetine showed better acceptability and tolerability but worse efficacy and response rate. In conclusion, most drugs are more effective than placebo, and there are few significant differences between the active drugs and placebo on acceptability. Overall, duloxetine and escitalopram showed better efficacy while vortioxetine showed better acceptability.

1. Introduction

Generalized anxiety disorder (GAD) is the most common anxiety disorder. GAD patients often present with excessive and persistent inexplicable tensions and fears with a lack of specific objects and specific content. This presentation is accompanied by symptoms of hyperactivity of somatic and autonomic nerves that severely affects their ability to work and study, and makes the patients experience pain and a lack of control (Craske and Stein, 2016). This disease has serious economic and social burdens globally.

Drugs are still the main clinical treatment method for GAD. The guide recommends SSRI and SNRI drugs as first-line treatments for

GAD, but there seems to be a lack of evidence for the clinician to choose which drug. Many head-to-head randomized controlled trials (RCTs) reported in the literature have investigated the efficacy of different drugs for the treatment of GAD, but the efficacy of these first-line drugs needs to be evaluated systematically (Gomez et al., 2018).

There are two articles which systematically reviewed the efficacy of anxiolytic drugs, which included all mechanistic drugs (Baldwin et al., 2011; Slee et al., 2019). Different from these two articles, we only compare the common first-line treatment drugs of GAD, which contained SSRIs (fluoxetine, paroxetine, sertraline, fluvoxamine, citalopram and escitalopram) and SNRIs (duloxetine and venlafaxine). In addition, some new drugs have subsequently been applied in the

* Corresponding author. Department of Psychiatry; Clinical Research center for Mental Disease of Shaanxi Province; Center for Brain Science, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, 710061, China.

** Corresponding author.

E-mail addresses: lujun2006@xjtu.edu.cn (J. Lyu), maxiancang@163.com (X. Ma).

<https://doi.org/10.1016/j.jpsychires.2019.08.009>

Received 5 July 2019; Received in revised form 13 August 2019; Accepted 16 August 2019

0022-3956/ © 2019 Elsevier Ltd. All rights reserved.

treatment of GAD: vilazodone in January 2011 (Zareifopoulos and Dylja, 2017), levomilnacipran in July 2013 (Blier et al., 2017), vortioxetine in September 2013 (Christensen et al., 2017). These drugs have other antidepressant mechanisms as well as the mechanisms of SSRIs or SNRIs, and their therapeutic effects may be significantly better than other drugs, which also needed to be assessed systematically.

The present study performed a network meta-analysis of fluoxetine, paroxetine, sertraline, fluvoxamine, citalopram, escitalopram, duloxetine, venlafaxine, vilazodone, levomilnacipran, and vortioxetine. All of the RCTs for treating GAD with one or more of these drugs were searched. The aim of this study was to comprehensively combine evidence from direct and indirect comparisons and identify the most advantageous drug for GAD from the perspectives of both efficacy and safety.

2. Methods

2.1. Study selection

The entire MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials, PsycINFO, SCI (Science Citation Index), ClinicalTrials.gov, Current Controlled Trials, World Health Organization IRIS (Institutional Repository for Information Sharing), and ISRCTN Registry databases were searched from 1980 up to January 1, 2019 to identify studies relevant to the efficacy and acceptability of 11 drugs in GAD treatment. We used both a free-text words search strategy and a Medical Subject Headings (MeSH) words search strategy. The free-text search terms were (“anxiety [ti ab]” OR “generalized adj anxiety adj disorder” OR “generalized adj anxiety adj disorder” OR “GAD” OR “HAM-A”) AND (“randomized controlled trial [pt]” OR “controlled clinical trial [pt]” OR “randomized [ti ab]” OR “placebo [ti ab]” OR “drug therapy [sh]” OR “randomly [ti ab]” OR “trial [ti ab]” OR “groups [ti ab]”) combined with a list of all included drugs (restricted to titles and Abstracts). The MeSH word search terms were (“anxiety disorder” AND (“clinical trial [Publication Type]” OR “clinical trials as topic” OR “controlled clinical trial [Publication Type]” OR “clinical trial, phase IV [Publication Type]” OR “clinical trial, phase III [Publication Type]” OR “clinical trial, phase II [Publication Type]” OR “pragmatic clinical trial [Publication Type]” OR “randomized controlled trial [Publication Type]” OR “pragmatic clinical trials as topic”)) combined with a list MeSH words of all the included drugs. The species was limited to humans and the language to English. The reference lists of identified systematic reviews, meta-analysis, and pooled analysis were reviewed to identify further studies.

2.2. Inclusion criteria

The process of article screening involved conducting initial screening based on topics and abstracts to select RCTs and reviews that may be relevant to the topic, downloading the full text for secondary screening, and then finally reading the review to check for missing articles.

Two researchers independently conducted article retrieval and screening. If both of them considered that an article did not meet the inclusion criteria, the article was excluded. If there was disagreement between these two people, they could either resolve this by negotiation or obtain the judgment of a third person. The screening was performed according to the literature inclusion criteria described below.

2.2.1. Type of study

Randomized parallel controlled studies were included, while crossover quasi-RCTs were excluded. The included RCTs must have researched a treatment for the acute phase of GAD, and literature researching relapses, changing to other drugs, or combination therapy was excluded.

2.2.2. Types of participants

Patients aged 18 years or older, of either sex, and with a primary diagnosis of GAD based on international diagnostic criteria such as Diagnostic and Statistical Manual of Mental Disorders (DSM)-III, DSM-IV, DSM-V, International Classification of Diseases (ICD)-9, and ICD-10 were included. A concurrent other psychiatric disorder was not considered an exclusion criterion. Studies including patients with other concurrent organic diseases, alcohol addiction or substance abuse were excluded. For a study including multiple types of anxiety disorder (e.g., panic disorder or social disorder) in the population, if information about the GAD population could be obtained, only the GAD population was included; in other situations the study was excluded.

2.2.3. Interventions

The intervention group included one or more of the following drugs: fluoxetine, paroxetine, sertraline, fluvoxamine, citalopram, escitalopram, duloxetine and venlafaxine, vilazodone, levomilnacipran, and vortioxetine. The control group was another of the above drugs or placebo. Studies that included only one drug (even at different dosages or with different formulations) were excluded.

2.2.4. Evaluation scale

The Hamilton Anxiety Scale (HAMA) was used to evaluate improvements in anxiety, and so literature that did not use this scale was rejected.

2.2.5. Length of RCT

The included RCTs needed to be ≥ 4 weeks long. We defined acute treatment as 8 weeks of treatment for both the efficacy and acceptability analyses. If 8-week data were not available, we used data from between 4 and 12 weeks, with preference given to the time point given in the original study as the study endpoint.

2.2.6. Analysis sample

For secondary research that was based on a sample that had also been used in one or more other RCTs, we only included the original RCT. The intention-to-treat (ITT) population was prioritized in the efficacy analysis. The ITT population was defined as receiving drug therapy and having at least one outcome assessment. If an ITT population was not reported on, the safety population was used as the second choice. The safety population was defined as the population treated with the medication. The safety population was the first choice for the acceptability analysis, and randomized populations were used if the safety population was not reported.

2.3. Data extraction

The following basic information was extracted from the included studies according to a table set up in advance using an electronic data capture system: first author, publication year, diagnostic criteria, sample size, age distribution, sex ratio, main inclusion criteria, registration information, sponsorship information, primary outcomes, length of RCT, analysis population, population setting, conducted sites, intervention drug, control drug, sample sizes of the intervention and control groups, changes in HAMA scores (mean \pm SD values), response rate, remission rate, safety population, the number of people who withdrew from the RCT (at 4–12 weeks) for any reason, and the number of people who withdrew from the RCT (at 4–12 weeks) due to adverse reactions.

2.4. Quality assessment

The methodological quality assessment of the included studies was based on the Cochrane Collaboration's tool in Review Manager Software, which comprises the following seven items: randomization method, concealment of allocation, blinding of outcome assessors, blinding of study personnel and participants, incomplete outcome data,

selective outcome reporting, and other sources of bias. The bias risk of each item was divided into three categories: low, unclear, and high. The study quality was divided into three levels from high to low: low-risk studies had no high-risk items and fewer than three items of unclear risk, medium-risk studies had one high-risk item or more than four items of unclear risk, and high-risk studies covered all other situations (Cipriani et al., 2018).

2.5. Outcomes

The primary outcomes of this meta-analysis were the efficacy and acceptability of the analyzed drugs. The efficacy measure was the change in HAMA score from baseline. If the mean difference was not provided, we attempted to obtain this using one of the following methods: calculating it from HAMA scores at baseline and post-treatment, asking the original author for the data, or estimating it from the figures in the article. If the SD was not provided, one of the following determination methods was chosen: asking the original author for the data; using the data in other meta-analysis; using the largest SD for the same medication. Acceptability was quantified based on treatment discontinuation, which was measured as the proportion of patients who withdrew for any reason from the safety population. Randomized populations were used if the safety population was not available.

The secondary outcomes were the response rate and tolerability. The response rate was measured as the total number of patients whose total score on HAMA reduced by $\geq 50\%$. The tolerability was measured as the proportion of patients who dropped out early from the safety population due to adverse events. Randomized populations were used if the safety population was not available.

2.6. Dosage selection

We included only study arms that randomized patients to drugs within the licensed dosage. Both fixed-dosage and flexible-dosage designs were allowed. A drug should be applied at the maximum dosage recommended in the usage guideline when assessing if it is clinically effective. Moreover, early increases in drug dosage can benefit patients earlier, and recent studies have shown that early improvement of symptoms may predict positive symptomatic outcomes (Habert et al., 2016; Kudlow et al., 2012). Therefore, when the same drug was administered in a study at different fixed dosages, the effect when it was applied at the maximum dosage in the recommended guideline was assessed. When a positive control drug was reported in the literature, a dosage comparable to that of the positive control drug was selected for the experimental drug. For example, when a positive control drug was applied at a high dosage, the comparison experimental group also applied it at a high dosage. The dosage used in efficacy analysis was the same as that used in other analyses of primary and secondary outcomes.

2.7. Data analysis

We first generated descriptive statistics for the trial design and the study population characteristics across all eligible trials, including comparative drug, clinical, and methodological characteristics such as the year of publication, age distribution of the population, severity of disease, sponsorship, and clinical settings.

A meta-analysis of pairwise comparisons was performed using STATA software. For each pairwise comparison, the weighted mean difference (WMD) was calculated as the effect size of the mean change from baseline of the HAMA score, and the odds ratio (OR) was calculated as the effect size for dichotomous outcomes such as acceptability, response rate, and tolerability. A probability value of $P < 0.05$ was considered to indicate a significant difference. A pooled analysis was conducted using a random-effects model with the assumption that different studies have different treatment effects. The presence of heterogeneity was assessed using Cochran's Q-statistic and quantified using

the I^2 statistic. A P value for Cochran's Q-statistic of < 0.1 or an I^2 value above 50% indicates the presence of a very high degree of heterogeneity.

The network meta-analysis was performed using the GeMTC package of R and JAGS (Just Another Gibbs Sampler) software in combination. The model calculations based on Bayesian theory were performed using the rjags package in JAGS, with $n. adapt = 5000$ and $n. iter = 20000$. Convergence was checked by running multiple chains and monitoring their mixing using the Brooks-Gelman-Rubin diagnostic. We also used that model to draw forest plots and ranking maps and perform consistency and heterogeneity tests.

We used STATA software to analyze inconsistency, publication bias and draw network diagrams. Consistency in network meta-analysis refers to the similarity between direct and indirect comparison results. So we compared results from the direct and indirect comparisons to analysis inconsistency. In addition, loop inconsistency refers to an inconsistency test based on a variety of measures that constitute a "ring", and is generally considered to be the basis for inconsistency. We therefore also analyze loop inconsistency (Higgins et al., 2012). Publication bias was evaluated using funnel plots. A comparison-adjusted funnel plot involves applying the classical funnel graph extension to network meta-analysis. It is necessary to calculate the study effect estimators for different comparisons in order to evaluate whether there are small-sample effects in the intervention network (Chaimani and Salanti, 2012). The horizontal axis of the graph indicates the difference between the effect of one paired comparison in a study and the combined effect of many similar comparisons, while the vertical axis normally shows the standard error of the effect size. Comparison-adjusted funnel plots will be symmetrical about the zero line if no small-sample effects are present.

3. Results

3.1. Study selection process and overview of the study characteristics

Fig. 1 shows the literature screening process. In total, 41 full articles met the inclusion criteria. The basic information for the 41 articles is listed in Table 1. These researches involved 15739 patients with GAD who were randomly assigned to receive 8 active drugs (duloxetine, venlafaxine, escitalopram, fluoxetine, paroxetine, sertraline, vilazodone and vortioxetine) or placebo. The mean study sample size was 384, the mean age of the patients was 41.83 years, and the proportion of women

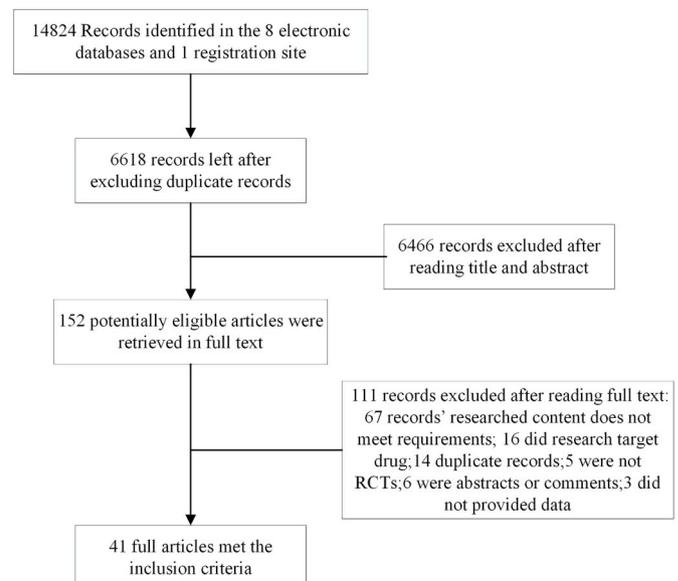


Fig. 1. Flow chat of study selection.

Table 1
Characteristic of included studies.

Study	Year	Length of RCT (week)	No. of patients	% female	Mean age (SD)	Population setting	Diagnostic criteria	Inclusion criteria	Multi/single center	recruitment	Patient numbers and every treatment groups	Funder
Qian ¹	2017	8	205	0.76	44.6	outpatient and inpatient	DSM-IV	NA	Single-center	China	104 ESCI 101 VENL	NA
Durgam ²	2016	8	404	0.65	39.9	outpatients	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; CGI-S ≥ 4; HAMDI7 ≤ 17	Multi-center	US	200 VILA 20–40 mg 200 PLAC	Forest Laboratories
Gommoll ³	2015	8	673	0.64	40.2	outpatients	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; CGI-S ≥ 4; HAMDI7 ≤ 17	Multi-center	US	223 VILA 20 mg 223 VILA 40 mg 221 PLAC	Forest Laboratories
Gommoll ⁴	2015	8	398	0.69	40	outpatients	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; CGI-S ≥ 4; HAMDI7 ≤ 17	Multi-center	US	198 VILA 20–40 mg 197 PLAC	Forest Laboratories
Alaka ⁵	2014	10	291	0.78	71.56 (5.22)	outpatients	DSM-IV TR	HAMD17 ≤ 17 CAS ≥ 9; no item of RDS > 3; CAS > RDS; HADS-anxiety ≥ 10	Multi-center	Cross-Continental	143 DULO 30–120 mg 131 PLAC	Eli Lilly and company
Mahableshwarkar ⁶	2014	8	781	0.68	38.6 (12.14)	outpatients and inpatients	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; MADRS ≤ 16	Multi-center	US	118 VORT 2.5 mg 114 VORT 5 mg 111 VORT 10 mg 106 DULO 60 mg 120 PLAC	Takeda Pharmaceutical Company, Ltd and H. Lundbeck A/S
Mahableshwarkar ⁷	2014	8	457	0.67	41.2 (14.15)	outpatient and inpatient	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; MADRS < 16	Multi-center	US	144 VORT 5 mg 146 VORT 10 mg 148 PLAC	Takeda Pharmaceutical Company, Ltd and H. Lundbeck A/S
Kaspe ⁸	2014	10	523	0.74	45.8	outpatients	DSM-IV-TR	HAMA ≥ 18; HAMA item 1 and 2 ≥ 2; HAMA points for psychic anxiety ≤ 21; CAS ≥ 9	Multi-center	Germany	132 PARO 20 mg 135 PLAC	Dr Willmar Schwabe GmbH & Co. KG
Stein ⁹	2014	12	412	0.72	42.6 (12.4)	outpatients	DSM-IV-TR	HAMA ≥ 22; HAMA item 1 and 2 ≥ 2; HAMA item 1 + item 2 > 5; MADRS ≤ 16; HADS-anxiety > HADS-depression	Multi-center	Cross-Continental	139 ESCI 10–20 mg 131 PLAC	Servier
Bidzan ¹⁰	2012	8	301	0.65	45.2	NA	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; MADRS ≤ 16	Multi-center	Cross-Continental	128 VORT 5 mg 126 PLAC	TakedaPharmac-eutical Company
Merideth ¹¹	2012	8	828	0.67	38.52	outpatients	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; MADRS ≤ 16; CGI-S ≥ 4	Multi-center	US	203 ESCI 10 mg 212 PLAC	AstraZeneca
Rothschild ¹²	2012	8	304	0.66	41.2	NA	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; MADRS ≤ 16	Multi-center	US	145 VORT 5 mg 144 PLAC	Takeda Pharmaceutical Company
Wu ¹³	2011	15	210	0.50	37.64	outpatients	DSM-IV	CGI-S ≥ 4; CAS > RDS; no item of RDS > 3; SDS ≥ 12	Multi-center	China	107 DULO 60–120 mg 100 PLAC	Eli Lilly and company

(continued on next page)

Table 1 (continued)

Study	Year	Length of RCT (week)	No. of patients	% female	Mean age (SD)	Population setting	Diagnostic criteria	Inclusion_criteria	Multi/single center	recruitment	Patient numbers and every treatment groups	Funder
Bandelow ¹⁴	2010	8	866	0.65	41.45	outpatients	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; MADRS ≤ 16; CGI-S ≥ 4	Multi-center	Cross-Continental	214 PARO 20 mg 217 PLAC	AstraZeneca Pharmaceuticals
Coric ¹⁵	2010	8	260	1	38.88	outpatients	DSM-IV-TR	HAMA ≥ 18; CGI-S ≥ 4	Multi-center	US	47 ESCI 20 mg 98 PLAC	Bristol-Myers Squibb
Feltner ¹⁶	2009	4	169	0.58	36.09	NA	DSM-IV	HAMA ≥ 20; RDS ≤ 7; CAS ≥ 9	Multi-center	multicenter	48 PARO 20 mg 54 PLAC	Pfizer
Nicolini ¹⁷	2009	10	581	0.57	42.8	outpatients	DSM-IV	HADS-anxiety ≥ 10; CAS ≥ 9; CAD > RDS none of the RDS items > 3; CGI-S ≥ 4	Multi-center	Cross-Continental	83 DULO 20 mg 151 DULO 60–120 mg 158 VENL 75–225 mg 163 PLAC	Eli Lilly and company
Kasper ¹⁸	2009	8	374	0.49	40.78	outpatients	DSM-IV-TR	HAMA ≥ 20; HAMA psychic and somatic anxiety factors ≥ 10	Multi-center	Cross-Continental	125 VENL 75–225 mg 128 PLAC	Pfizer Inc.
Rynn ¹⁹	2008	10	327	0.62	41.6	outpatients	DSM-IV	CGI-S ≥ 4; HADS-anxiety ≥ 10; CAS ≥ 9; CAS > RDS; no item of RDS > 3	Multi-center	US	168 DULO 60–120 mg 159 PLAC	Eli Lilly and Company
Bose ²⁰	2008	8	393	0.62	37.6	Outpatients	DSM-IV	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2	Multi-center	US	125 ESCI10-20 mg 125VENL75-225 mg 135PLAC	Forest Laboratories
Hartford ²¹	2007	10	487	0.63	40.8	outpatients	DSM-IV	HAMD ≤ 15; HADS-anxiety ≥ 10; CAS ≥ 9; no item of RDS > 3; CAS > RDS; CGI-S ≥ 4	Multi-center	US	162 DULO60-120 mg 164 VENL 75–225 mg 161 PLAC	Eli Lilly and Company and Boehringer Ingelheim
Koponen ²²	2007	9	513	0.68	43.8 (13)	outpatients	DSM-IV	CGI-S ≥ 4; HADS-anxiety ≥ 10; CAS ≥ 9; CAS > RDS; no item of RDS > 3	Multi-center	Cross-Continental	168 DULO 60 mg 170 DULO 120 mg 175 PLAC	Eli Lilly and Co. and Boehringer Ingelheim
Baldwin ²³	2006	12	682	0.64	41	outpatients	DSM-IV-TR	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; MADRS ≤ 16	Multi-center	Europe	134 ESCI 5 mg 134 ESCI 10 mg 132 ESCI 20 mg 138 PLAC	Lundbeck and GlaxoSmithKline.
Brawman-Mintzer ²⁴	2006	10	326	0.58	40.45	outpatients	DSM-IV	HAMA ≥ 20; HAMA item 1 ≥ 2; CAS > RDS	Multi-center	US	136 PARO 164 SERT 50–200 mg 162 PLAC	Pfizer Inc.
Kim ²⁵	2006	8	46	0.61	44.35	outpatients	DSM-IV	HAMA ≥ 18; CGI-S ≥ 4	Single-center	Korea	21 VENL 37.5–220 mg 25 PARO 10–40 mg	None
Montgomery ²⁶	2006	6	421	0.62	41.4 (12.3)	outpatients	DSM-IV	HAMA ≥ 20; CAS ≥ 9; RDS ≤ 7	Multi-center	Europe	110 VENL 75 mg 100 PLAC	Pfizer Inc.
Ball ²⁷	2005	8	53	0.77	39.46	NA	DSM-IV	HAMA ≥ 18; HAMD ≤ 20	Single-center	indian	25 PARO 10–40 mg 28 SERT 25–100 mg	Pfizer Inc.
Bielski ²⁸	2005	24	123	0.62	37	none	DSM-IV	HAMA ≥ 18; HAMD ≥ 17; CAS > RDS	Multi-center	US	60 ESCI 10–20 mg 61 PARO 20–50 mg	Forest Laboratories, Inc.
Allgulander ²⁹	2004	12	373	0.55	41.36	outpatients	DSM-IV	HAMA ≥ 18; HAMA item 1 and 2 ≥ 2	Multi-center	Cross-Continental	182 SERT 50–150 mg 188 PLAC	Pfizer, Inc.
Davidson ³⁰	2004	8	315	0.53	39.5	outpatients	DSM-IV	HAMA ≥ 18; HAMA item 1 + 2 ≥ 5	NA	NA	154 ESCI 10–20 mg 153 PLAC	Forest Laboratories, Inc.

(continued on next page)

Table 1 (continued)

Study	Year	Length of RCT (week)	No. of patients	% female	Mean age (SD)	Population setting	Diagnostic criteria	Inclusion_criteria	Multi/single center	recruitment	Patient numbers and every treatment groups	Funder
Nimatoudis ³¹	2004	8	46	0.65	42.43	outpatients	DSM-IV	HAMD < 17; CAS > RDS HAMA ≥ 18; CAS ≥ 8	Multi-center	Greece	24 VENL 75–150 mg 22 PLAC	None
Lenox-Smith ³²	2003	24	244	0.59	47	outpatients	DSM-IV	HAMA ≥ 20; MADRS < 23 CAS ≥ 8	Multi-center	UK	122 VENL 75 mg 122 PLAC	Wyeth
Rickels ³³	2003	8	566	0.55	40.5	outpatients	DSM-IV	HAMA ≥ 22; HAMA item 1 and 2 ≥ 2; MADRS ≤ 17	Multi-center	North America	188 PARO 20 mg 197 PARO 40 mg 180 PLAC	GlaxoSmithKline
Hackett ³⁴	2003	8	556	0.66	44.2	outpatients	DSM-IV	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2 CAS > RDS RDS ≤ 9	Multi-center	NA	185 VENL 75mg 169 VENL 150 mg 97 PLAC	Wyeth
Allgulander ³⁵	2001	24	541	0.61	45.14	outpatients	DSM-IV	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2	Multi-center	Europe	138 VENL 37.5 mg 130 VENL 75 mg 131 VENL 150 mg 130 PLAC	Wyeth-Ayerst Research
Pollack ³⁶	2001	8	324	0.80	40.5	outpatients	DSM-IV	HAMA ≥ 20; HAMA item 1 and 2 ≥ 2; MADRS < 17	Multi-center	North America	161 PARO 20–50 mg 163 PLAC	GlaxoSmithKline, Philadelphia, Pa
Silverstone ³⁷	2001	12	90	0.69	43.5 (11.1)	NA	DSM-IV	NA	Multi-center	NA	32 VENL 75-225 mg 33 FLUO 20–60 mg 25 PLAC	Wyeth-Ayerst research
Gelenberg ³⁸	2000	28	238	0.59	36.45	outpatients	DSM-IV	HAMA ≥ 18; HAMA item 1 and 2 ≥ 2 RDS ≥ 9; CAS > RDS	Multi-center	US	115 VENL 75–225 mg 123 PLAC	Wyeth-Ayerst Research, Philadelphia, Pa
McCafferty ³⁹	2000	8	324	0.64	41.01	patinets	DSM-IV	HAMA ≥ 20 MADRS < 18 CAS > RDS	NA	NA	161 PARO 20–50 mg 163 PLAC	None
Rickels ⁴⁰	2000	8	349	0.56	40.84	outpatients	DSM-IV	HAMA ≥ 18; HAMA item1 and 2 ≥ 2 RDS ≤ 9; CAS > RDS	Multi-center	US	86 VENL 75 mg 81 VENL 150 mg 86 NEML 225 mg 96 PLAC	wyeth-Ayerst Laboratories
Davidson ⁴¹	1999	8	365	0.61	37.78	outpatients	DSM-IV	HAMA ≥ 18; HAMA item 1 and 2 ≥ 2 RDS ≤ 9; CAS > RDS; no item of RDS > 3	Multi-center	US	87 VENL 75 mg 87 VENL 150 mg 98 PLAC	Wyeth-Ayerst Research, Philadelphia, Pa

RCT, Random Controlled Trial; SD, Standard Deviation; HAMA, Hamilton Anxiety scale; MADRS, Montgomery-Åsberg Depression Rating Scale; CGI, Clinical Global Impressions-Severity; HAMD, Hamilton Depression scale; US, the United States; CAS, Covi Anxiety Scale; RDS, Raskin Depression Scale; HADS, Hospital Anxiety and Depression Scale; SDS, Sheehan Disability Scale; GDS, Geriatric Depression Scale; MMSE, Mini-Mental State Examination; PLAC, placebo; DULO, duloxetine; ESCI, Escitalopram; FLUO, Fluoxetine; PARO, paroxetine; SERT, sertraline; VENL, venlafaxine; VILA, Vilazodone; VORT, vortioxetine.

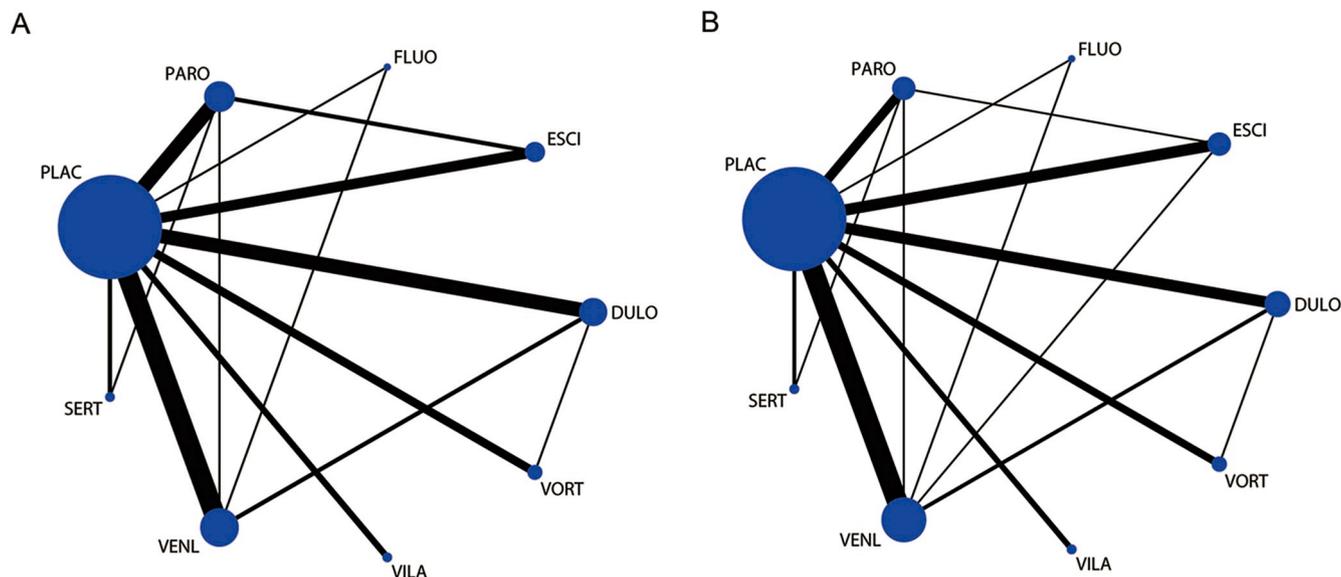


Fig. 2. Network meta-analysis of eligible comparisons for efficacy (A) and acceptability (B).

Width of the lines is proportional to the number of trials comparing every pair of treatments. Size of every circle is proportional to the number of randomly assigned participants. PLAC, placebo; DULO, duloxetine; ESCI, escitalopram; FLUO, fluoxetine; PARO, paroxetine; SERT, sertraline; VENL, venlafaxine; VILA, vilazodone; VORT, vortioxetine.

was 64.4%. GAD was diagnosed based on DSM-IV or DSM-IV-TR. Most (87.8%) had a multicenter design. The response rate was 49.4%. The proportion of patients who withdrew for any reason from the safety population was 25%, while 8.3% dropped out early due to adverse events. Pharmaceutical companies funded 37 (90.2%) of the 41 studies. [Supplementary Fig. S1](#) showed the result of the quality assessment. The risk of bias was high, moderate, and low in 15 (36.6%), 7 (17.1%), and 19 (46.3%) of the 41 RCTs, respectively.

3.2. Network meta-analysis

[Fig. 2](#) shows the network of eligible comparisons for efficacy and acceptability. Each of the active drugs was included in at least one placebo-controlled trial. Vilazodone was not directly compared with another active drug in the network. [Supplementary Table S1](#) provides detailed results of pairwise meta-analyses of primary outcomes. Except vortioxetine, all drugs show better efficacy than placebo. While for acceptability, vilazodone and paroxetine were worse than placebo, others were not significantly different from placebo. [Fig. 3](#) shows the results of the network meta-analysis for the primary and secondary outcomes. In terms of efficacy (covering 37 RCTs involving 10302 patients), all of the drugs except fluoxetine and vortioxetine were more effective than placebo, with the weighted mean difference ranging between -3.2 (95% credible interval [CrI] = -4.2 to -2.2) for escitalopram and -1.8 (95% CrI = -3.1 to -0.55) for vilazodone. For acceptability (covering 34 RCTs involving 10342 patients), only vilazodone (OR = 1.7, 95% CrI = 1.1 to 2.7) were worse than placebo, others did not show significant differences from placebo.

For the response rate (covering 33 RCTs involving 9284 patients), all of the drugs except fluoxetine and vortioxetine showing a higher response rate than placebo, with OR ranging between 2.2 (95% CrI = 1.8 to 2.8) for venlafaxine and 1.6 (95% CrI = 1.1 to 2.4) for paroxetine. In terms of dropouts due to adverse events, all of the drugs except vortioxetine, sertraline and fluoxetine were associated with higher withdrawal rates than placebo, with OR ranging between 3.8 (95% CrI = 2.1 to 7.3) for vilazodone and 1.9 (95% CrI = 1.3 to 3.1) for escitalopram.

3.3. Head-to-head comparisons

We also performed head-to-head comparisons of the primary outcome for the 8 active drugs; the results are shown in [Table 2](#). For efficacy, venlafaxine, escitalopram and duloxetine were more effective than vortioxetine (WMD ranging between -2.2 and -1.8). For acceptability, vortioxetine (OR = 0.51, 95% CrI = 0.29 to 0.9) and escitalopram (OR = 0.58, 95% CrI = 0.34 to 0.98) showed better acceptability than vilazodone.

3.4. Inconsistency

The results from the direct and indirect comparisons of efficacy were consistent for sertraline vs. placebo, escitalopram vs. placebo, paroxetine vs. placebo, paroxetine vs. sertraline, duloxetine vs. venlafaxine, paroxetine vs. venlafaxine, duloxetine vs. vortioxetine, paroxetine vs. escitalopram and venlafaxine vs. placebo. No inconsistency was found. Six closed loops were produced, and there was no inconsistent loop.

The results from the direct and indirect comparisons of acceptability were consistent for sertraline vs. placebo, venlafaxine vs. placebo, paroxetine vs. placebo, paroxetine vs. sertraline, duloxetine vs. venlafaxine, paroxetine vs. venlafaxine, duloxetine vs. vortioxetine, venlafaxine vs. escitalopram and paroxetine vs. escitalopram. No inconsistency was found. Eight loops were produced, and there was no inconsistent loop.

Regarding the secondary outcomes, the results from the direct and indirect comparisons of the response rate and tolerability were consistent and there are no inconsistent loop ([Supplementary Fig. S2](#)).

3.5. Publication bias

We used funnel and comparison-adjusted funnel plots to evaluate publication bias and the small-sample effect. The funnel plots and comparison-adjusted funnel plots for efficacy and acceptability are shown in [Fig. 4](#). For efficacy, it can be seen from the figure that the funnel plot is roughly symmetrical, and there do not appear to be any studies with small samples on the left-hand side of the zero line in the comparison-adjusted funnel plot, indicating the absence of obvious publication bias; however, a small-sample effect might still have been

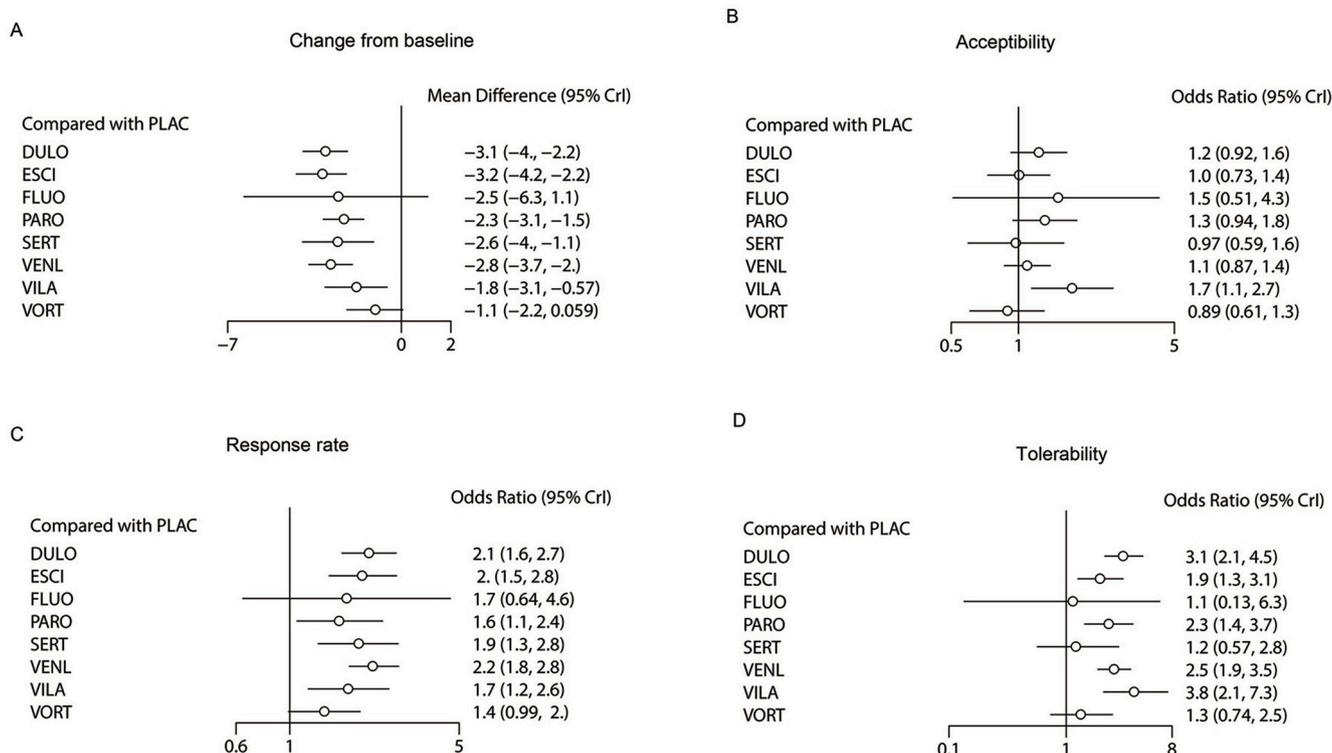


Fig. 3. Forest plots of network meta-analysis of all trials for efficacy (A), acceptability (B), response rate (C) and tolerability (D). Antianxiety drugs were compared with placebo, which was the reference compound. OR = odds ratio. CrI = credible interval. PLAC, placebo; DULO, duloxetine; ESCI, escitalopram; FLUO, fluoxetine; PARO, paroxetine; SERT, sertraline; VENL, venlafaxine; VILA, vilazodone; VORT, vortioxetine.

Table 2
Results of head-to-head comparison.

	VILA	DULO	VORT	PARO	ESCI	VENL	SERT	FLUO
VILA		0.71 (0.43–1.2)	0.51(0.29–0.9)	0.75 (0.44–1.3)	0.58(0.34–0.98)	0.63 (0.39–1)	0.56 (0.29–1.1)	0.87 (0.27–2.6)
DULO	-1.2 (-2.8–0.28)		0.72 (0.46–1.1)	1.1 (0.69–1.7)	0.82 (0.53–1.3)	0.88 (0.63–1.2)	0.79 (0.45–1.4)	1.2 (0.4–3.6)
VORT	0.77 (-0.94–2.5)	2(0.64–3.4)		1.5 (0.89–2.4)	1.1 (0.69–1.8)	1.2 (0.79–1.9)	1.1 (0.59–2.1)	1.7 (0.54–5.2)
PARO	-0.51 (-2–0.96)	0.74 (-0.49–1.9)	-1.3 (-2.7–0.15)		0.77 (0.5–1.2)	0.83 (0.56–1.2)	0.74 (0.42–1.3)	1.1 (0.36–3.4)
ESCI	-1.4 (-3–0.19)	-0.13 (-1.5–1.2)	-2.2(-3.7–0.64)	-0.87 (-2.1–0.26)		1.1 (0.74–1.6)	0.96 (0.54–1.7)	1.5 (0.48–4.5)
VENL	-1 (-2.6–0.46)	0.23 (0.93–1.4)	-1.8(-3.2–0.37)	-0.51 (-1.6–0.59)	0.36 (-0.92–1.7)		0.89 (0.51–1.5)	1.4 (0.46–3.9)
SERT	-0.74 (-2.7–1.1)	0.5 (-1.2–2.2)	-1.5 (-3.4–0.33)	-0.24 (-1.8–1.4)	0.64 (-1.1–2.4)	0.28 (-1.4–1.9)		1.5 (0.47–4.9)
FLUO	-0.89 (-4.8–3)	0.35 (-3.4–4.2)	-1.7 (-5.5–2.2)	-0.39 (-4.2–3.4)	0.5 (-3.3–4.4)	0.13 (-3.5–3.9)	-0.17 (-4.1–3.8)	

DULO, duloxetine; ESCI, Escitalopram; FLUO, Fluoxetine; PARO, paroxetine; SERT, sertraline; VENL, venlafaxine; VILA, Vilazodone; VORT, vortioxetine.

present. No obvious publication bias or small-sample effect was found for acceptability. The funnel plots for response rate and tolerability are shown in [Supplementary Fig. S3](#). There was no obvious publication bias, but a small-sample effect was present for the response rate.

4. Discussion

The present study found that most drugs show superior efficacy to placebo but without an improvement in acceptability. In addition, the head-to-head comparisons revealed that venlafaxine, escitalopram and duloxetine were more efficacious than vortioxetine. Vortioxetine and escitalopram showed better acceptability than vilazodone. We also found that sertraline it was well tolerated although it was not as effective as venlafaxine, escitalopram or duloxetine. These results may be help to inform patients, physicians, guideline developers, and policy makers on the relative merits of the different first-line treatment drug on GAD.

The effect size for efficacy was generally negative since it was calculated as the mean change in the HAMA total score determined as the final value after treatment minus the baseline value. The effect size for

acceptability/tolerability was the ratio of the number of people in the safety population who dropped out due to various reasons/adverse reactions, and a smaller value represents better acceptability/tolerability. This study strictly limited the inclusion criteria and ensured transferability between comparisons. In addition, consistency was the primary method for identifying the accuracy and reliability of the network meta-analysis (Higgins et al., 2012). The academic consensus is that all external interference factors in a network meta-analysis will impact the results obtained in direct and indirect comparisons. Therefore, unlike a traditional meta-analysis, consistency in a network meta-analysis refers to the similarity between the results from direct and indirect comparisons. This study found no inconsistency in the results from the direct and indirect comparisons, and no ring inconsistency was found in all of outcomes. Overall, the consistency of the present network meta-analysis was good and the results are reliable.

Most clinical guidelines recommend selective serotonin-reuptake inhibitors SSRIs and serotonin/norepinephrine-reuptake inhibitors (SNRIs) as the first-choice drugs for GAD. Among them, there is strong evidence for the use of escitalopram and venlafaxine (Canadian Psychiatric Association, 2006; Katzman et al., 2014). The result of this

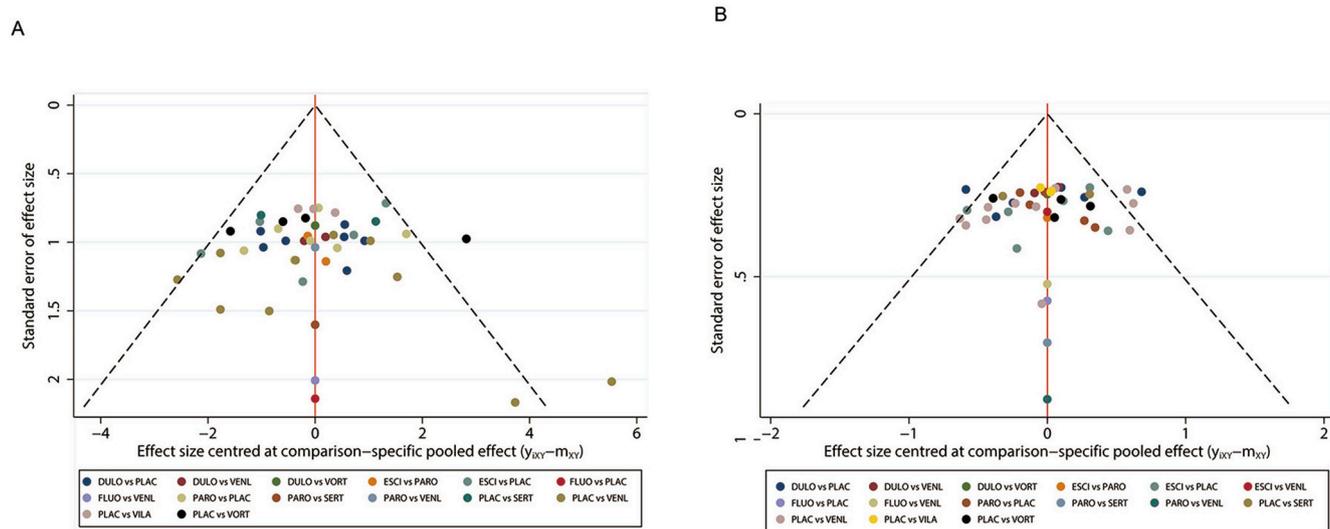


Fig. 4. The funnel plots and comparison-adjusted funnel plots for efficacy (A) and acceptability (B).

PLAC, placebo; DULO, duloxetine; ESCI, escitalopram; FLUO, fluoxetine; PARO, paroxetine; SERT, sertraline; VENL, venlafaxine; VILA, vilazodone; VORT, vortioxetine.

network meta-analysis are consistent with the results of the literature that escitalopram and venlafaxine performed well in terms of efficacy, with moderate acceptability. It is worth noting that we found that duloxetine also has a good effect, but its acceptability is not as good as venlafaxine and escitalopram.

A lot of literature found that the new drug vortioxetine has a good effect on the treatment of MDD. Our previous article also found that 10 mg/day vortioxetine might be optimal for the treatment of MDD (He et al., 2018). Because the mechanism of action of this drug includes SSRIs' action mechanism, it may also be effective for GAD. Unfortunately, the results of the current studies are inconsistent. Bidzan et al. found that the 5 mg dose of vortioxetine was more effective and safe in treating GAD than placebo (Bidzan et al., 2012), but some studies have found that any dose of vortioxetine does not differ from placebo on efficacy in the treatment of GAD, although its tolerance may be better (Mahableshwarkar et al., 2014a,b; Mahableshwarkar et al., 2014a,b; Rothschild et al., 2012). Our results are similar in the latter studies. In addition, Yee et al. found that vortioxetine has a better therapeutic effect and tolerance to total anxiety disorder (Yee et al., 2018), plus the results of vortioxetine in the treatment of other types of anxiety (Liebowitz et al., 2017; Shah and Northcutt, 2018), we speculated that this drug may have a certain effect on social phobia or panic disorder.

There are several possible causes for the heterogeneity present in the available research results. Firstly, this study did not exclude patients with other psychiatric diseases comorbid with GAD. Comorbidity is common in GAD, particularly depression and other anxiety disorders (Sandelin et al., 2013), and so excluding these patients would eliminate most of the literature and not meet the clinical inclusion criteria. However, it is not possible to examine the effects of comorbid depression or other anxiety disorders on clinical outcomes. Secondly, the severity of GAD patients in the included studies would have differed somewhat (HAMA scores greater than 18–22). It has been shown previously that the severity of GAD affects the efficacy of drug treatment (Butler, 1993). Thirdly, the drug dosages have not been consistent. These factors should be fully considered when interpreting the results of this study.

One limitation of this study is that unpublished data and newest data were not included and analyzed. Secondly, early clinical trials often have small samples, which can easily result in small-sample effects. Certain small-sample effects were present in our analyses of the efficacy, response rate, and tolerability. Due to the small number of

studies, this meta-analysis did not include a sensitivity analysis of small-sample studies, which could have introduced bias, while another limitation is the missing data. Although we conservatively estimated the missing data (especially the SD values), the results might still have been biased. In addition, most of the studies were funded by pharmaceutical companies, and hence the results may have been affected by funding bias. We did not attempt to overcome this problem by including unpublished studies, since we consider that they are more likely to have design or data-reliability issues.

The results of this study indicate that common SSRIs and SNRIs can be chosen to treat generalized anxiety disorder. There are differences in the efficacy and acceptability profiles across these drugs, and the clinician needs to choose the best drug according to the patient's specific situation. Our finding may contribute a helpful perspective on making these decisions.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 81771471), the National Natural Science Foundation of China (No.81460560) and the Clinical Research Award of the First Affiliated Hospital of Xi'an Jiaotong University, China (No.XJTU1AF-CRF-2016-024).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jpsychires.2019.08.009>.

References

- Baldwin, D., Woods, R., Lawson, R., Taylor, D., 2011. Efficacy of drug treatments for generalised anxiety disorder: systematic review and meta-analysis. *BMJ* 342, d1199.
- Bidzan, L., Mahableshwarkar, A.R., Jacobsen, P., Yan, M., Sheehan, D.V., 2012. Vortioxetine (Lu AA21004) in generalized anxiety disorder: results of an 8-week, multinational, randomized, double-blind, placebo-controlled clinical trial. *Eur. Neuropsychopharmacol.* 22, 847–857.
- Blier, P., Gommoll, C., Chen, C., Kramer, K., 2017. Effects of levomilnacipran ER on noradrenergic symptoms, anxiety symptoms, and functional impairment in adults with major depressive disorder: post hoc analysis of 5 clinical trials. *J. Affect. Disord.*

- 210, 273–279.
- Butler, G., 1993. Predicting outcome after treatment for generalised anxiety disorder. *Behav. Res. Ther.* 31, 211–213.
- Canadian Psychiatric Association, 2006. Clinical practice guidelines. Management of anxiety disorders. *Can. J. Psychiatr.* 51, 9S–91S.
- Chaimani, A., Salanti, G., 2012. Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. *Res. Synth. Methods* 3, 161–176.
- Christensen, M.C., Loft, H., Florea, I., McIntyre, R.S., 2017. Efficacy of vortioxetine in working patients with generalized anxiety disorder. *CNS Spectr.* 1–9.
- Cipriani, A., Furukawa, T.A., Salanti, G., Chaimani, A., Atkinson, L.Z., Ogawa, Y., Leucht, S., Ruhe, H.G., Turner, E.H., Higgins, J., Egger, M., Takeshima, N., Hayasaka, Y., Imai, H., Shinohara, K., Tajika, A., Ioannidis, J., Geddes, J.R., 2018. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 391, 1357–1366.
- Craske, M.G., Stein, M.B., 2016. Anxiety. *Lancet* 388, 3048–3059.
- Gomez, A.F., Barthel, A.L., Hofmann, S.G., 2018. Comparing the efficacy of benzodiazepines and serotonergic anti-depressants for adults with generalized anxiety disorder: a meta-analytic review. *Expert Opin. Pharmacother.* 19, 883–894.
- Habert, J., Katzman, M.A., Oluboka, O.J., McIntyre, R.S., McIntosh, D., MacQueen, G.M., Khullar, A., Milev, R.V., Kjemistad, K.D., Chokka, P.R., Kennedy, S.H., 2016. Functional recovery in major depressive disorder: focus on early optimized treatment. *Prim. Care. Comp. CNS. Disord.* 18.
- He, H., Wang, W., Lyu, J., Zheng, J., Guo, L., An, X., Fan, Y., Ma, X., 2018. Efficacy and tolerability of different doses of three new antidepressants for treating major depressive disorder: a PRISMA-compliant meta-analysis. *J. Psychiatr. Res.* 96, 247–259.
- Higgins, J.P., Jackson, D., Barrett, J.K., Lu, G., Ades, A.E., White, I.R., 2012. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res. Synth. Methods* 3, 98–110.
- Katzman, M.A., Bleau, P., Blier, P., Chokka, P., Kjemistad, K., Van Ameringen, M., Antony, M.M., Bouchard, S., Brunet, A., Flament, M., Grigoriadis, S., Mendlowitz, S., O'Connor, K., Rabheru, K., Richter, P.M., Robichaud, M., Walker, J.R., 2014. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. *BMC Psychiatry* 14 (Suppl. 1), S1.
- Kudlow, P.A., Cha, D.S., McIntyre, R.S., 2012. Predicting treatment response in major depressive disorder: the impact of early symptomatic improvement. *Can. J. Psychiatr.* 57, 782–788.
- Liebowitz, M.R., Careri, J., Blatt, K., Draine, A., Morita, J., Moran, M., Hanover, R., 2017. Vortioxetine versus placebo in major depressive disorder comorbid with social anxiety disorder. *Depress. Anxiety* 34, 1164–1172.
- Mahableshwarkar, A.R., Jacobsen, P.L., Chen, Y., Simon, J.S., 2014a. A randomised, double-blind, placebo-controlled, duloxetine-referenced study of the efficacy and tolerability of vortioxetine in the acute treatment of adults with generalised anxiety disorder. *Int. J. Clin. Pract.* 68, 49–59.
- Mahableshwarkar, A.R., Jacobsen, P.L., Serenko, M., Chen, Y., 2014b. A randomized, double-blind, fixed-dose study comparing the efficacy and tolerability of vortioxetine 2.5 and 10 mg in acute treatment of adults with generalized anxiety disorder. *Hum. Psychopharmacol.* 29, 64–72.
- Rothschild, A.J., Mahableshwarkar, A.R., Jacobsen, P., Yan, M., Sheehan, D.V., 2012. Vortioxetine (Lu AA21004) 5 mg in generalized anxiety disorder: results of an 8-week randomized, double-blind, placebo-controlled clinical trial in the United States. *Eur. Neuropsychopharmacol.* 22, 858–866.
- Sandelin, R., Kowalski, J., Ahnemark, E., Allgulander, C., 2013. Treatment patterns and costs in patients with generalised anxiety disorder: one-year retrospective analysis of data from national registers in Sweden. *Eur. Psychiatry* 28, 125–133.
- Shah, A., Northcutt, J., 2018. An open-label, flexible dose adaptive study evaluating the efficacy of vortioxetine in subjects with panic disorder. *Ann. Gen. Psychiatr.* 17, 19.
- Slee, A., Nazareth, I., Bondaronek, P., Liu, Y., Cheng, Z., Freemantle, N., 2019. Pharmacological treatments for generalised anxiety disorder: a systematic review and network meta-analysis. *Lancet* 393, 768–777.
- Yee, A., Ng, C.G., Seng, L.H., 2018. Vortioxetine treatment for anxiety disorder: a meta-analysis study. *Curr. Drug Targets* 19, 1412–1423.
- Zareifopoulos, N., Dylja, I., 2017. Efficacy and tolerability of vilazodone for the acute treatment of generalized anxiety disorder: a meta-analysis. *Asian. J. Psychiatr.* 26, 115–122.