Psychopharm

Journal of Psychopharmacology
21(8) (2007) 864–872

ISSN 0269-8811
SAGE Publications
Los Angeles, London,
New Delhi and Singapore
10.1177/0269881107076996

An effect-size analysis of pharmacologic treatments for generalized anxiety disorder

Rosario B. Hidalgo *Anxiety and Traumatic Stress Program, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA.*

Larry A. Tupler *Anxiety and Traumatic Stress Program, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA.*

Jonathan R. T. Davidson *Anxiety and Traumatic Stress Program, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA.*

Abstract

Generalized anxiety disorder (GAD) is a prevalent and impairing disorder, associated with extensive psychiatric and medical comorbidity and usually characterized by a chronic course. Different drugs have been investigated in GAD; among them are the following: 1) SSRIs: paroxetine, sertraline, fluvoxamine and escitalopram; 2) SNRI1s: venlafaxine; 3) benzodiazepines (BZs): alprazolam, diazepam and lorazepam; 4) azapirones (AZAs): buspirone; 5) antihistamines (AHs): hydroxyzine; 6) pregabalin (PGB); and 7) complementary/alternative medicine (CAM): kava-kava and homeopathic preparation. We conducted an effect size (ES) analysis of 21 double-blind placebo-controlled trials of medications treating DSM-III-R, DSM-IV or ICD-10 GAD using HAM-A change in score from baseline or endpoint score as the main efficacy measure. Literature search was performed using MEDLINE and PsycINFO databases including articles published between 1987 and 2003 and personal communications with investigators and sponsors. Comparing all drugs versus placebo, the ES was 0.39. Mean ESs, excluding children, were PGB: 0.50, AH: 0.45, SNRI: 0.42, BZ: 0.38, SSRI: 0.36, AZA: 0.17 and CAM: −0.31. Comparing ES for adults versus children/adolescents (excluding CAM) and conventional drugs versus CAM (excluding children/adolescents) we found significantly higher ES for children/adolescents and for conventional drugs ($p < 0.001$ and $p < 0.01$, respectively). No significant differences were found when comparing date of publication, location of site (i.e. US versus other), fixed versus flexible dosing, number of study arms, or number of outcome measures used. Medications varied in the magnitude of their ES, ranging from moderate to poor. Adolescents and children showed a much greater ES compared with adults. Subjects taking CAM had worse outcomes than placebo.

Keywords

GAD, pharmacologic treatment, effect size, meta-analysis

Introduction

Generalized anxiety disorder (GAD) has a lifetime prevalence of 5.1% in the general population (Wittchen *et al*., 1994), is associated with extensive psychiatric and medical comorbidity (Hidalgo and Davidson, 2001) and is usually characterized by a chronic course.

Various medications have been investigated and demonstrated a range of degrees of efficacy in the treatment of GAD; among them are: 1) selective serotonin reuptake inhibitors (SSRIs): paroxetine, sertraline, escitalopram and fluvoxamine; 2) serotonin noradrenaline reuptake inhibitors (SNRIs): venlafaxine; 3) benzodiazepines (BZs): alprazolam, lorazepam and diazepam; 4) azapirones (AZAs): buspirone; 5) antihistamines (AHs): hydroxyzine; 6) pregabalin (PGB; an α-2-delta subunit calcium channel blocker) and 7) complementary/alternative medicines (CAMs): kava-kava and homeopathic preparation.

To date, three antidepressants have been approved by the Food and Drug Administration (FDA) for the treatment of GAD: escitalopram, paroxetine and venlafaxine, but a larger database of effective non-approved treatments also exists.

Our purpose is to review the available data from drug trials for the treatment of GAD as defined by DSM-III-R, DSM-IV or ICD-10 and to report a meta-analysis of their findings examining (1) the efficacy of the different compounds in terms of effect size (ES) compared with placebo and (2) how different variables may influence treatment response. Specifically, and partly influenced by the findings from the depression literature, we were interested in assessing the possible influence of the following variables: flexible

Corresponding author: Jonathan R. T. Davidson, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Box 3812, Durham, NC 27710, USA. *Email:* david011@mc.duke.edu

versus fixed dose (Khan *et al*., 2003); duration of treatment (Khan *et al*., 2000); numbers of measures used; number of treatment arms (Zimmerman and Posternak, 2003); year of publication (Walsh *et al*., 2002); geographic location (North America [US and Canada] versus outside North America [Europe/Australia]); adults versus children and adolescents; conventional medicines versus CAM (Pittler and Ernst, 2000).

We also examined possible publication bias to assess the robustness of our meta-analysis by computing the fail-safe *N*.

Given recent controversies regarding the use of medications in children and adolescents, we included two studies focusing on this population group (Rynn *et al*., 2001; Walkup *et al*., 2001). The use of CAM by the general population has grown in recent years, and for that reason we included two available trials in our analysis. We also included two trials of PGB, a drug which has been studied extensively in GAD but which is marketed in the United States for treatment of pain associated with diabetic neuropathy and post-herpetic neuralgia.

Materials and methods

Double-blind, placebo-controlled studies of different medications for treatment of GAD were reviewed through a literature search in MEDLINE and PsychINFO. We also included personal communications with investigators and sponsors of studies that met inclusion criteria but had not been published by the time of data collection. The following sponsors or investigators were contacted to provide data which we could not retrieve from the publication or from studies that had not yet been published: Forest Laboratories (Andrew Korotzer, escitalopram data); GlaxoSmithKline (Stan Krulewicz, paroxetine data); NIH/NIMH (Benedetto Vitiello, subgroup with GAD in the children/adolescents fluvoxamine data); Pfizer (data on sertraline); Wyeth (David Hackett, M.D., venlafaxine XR data). We did not attempt to contact the investigators of the hydroxizine study for which bromazepam data was not provided (Llorca *et al*., 2002). Each one of the sponsors and/or the investigators contacted were generous to provide the information requested. In some cases (e.g. Walkup *et al*., 2001; NIH/NIMH, Benedetto Vitiello personal communication November, 2003; Forest Laboratories, Andrew Korotzer personal communication, January–June, 2004) we obtained subgroup analysis that had not been presented in the original report.

Key words used in our literature search, alone and in combination, were: generalized anxiety disorder, treatment outcome, medication, SSRI, SNRI, benzodiazepines and the individual name of different drugs.

Because the concept of GAD has changed over the years (Rickels and Rynn, 2001) we decided to include studies that used DSM-III-R, DSM-IV and ICD-10 criteria for GAD; i.e. we assessed GAD as a chronic disorder, in contrast to DSM-III or previous versions, which included briefer episodes and are now outdated. For the same reason, we limited our search to articles published from 1987 to 2003, taking into account that prior to 1987 DSM-III criteria were likely to be used. However, a recent review found no difference between DSM-III, -III-R and -IV GAD with respect to effect size (Mitte *et al*., 2005).

We required at least two studies examining a particular class of drug (e.g. SSRI, CAM). Only English-language publications were included. We excluded studies presenting uncontrolled trials, case reports, reviews of trials that were published separately and studies comparing medication with psychotherapy. A number of reports (e.g. concerning abecarnil) were excluded because of insufficient information in the publication to allow calculation.

We focused on the Hamilton Rating Scale for Anxiety (HAM-A) (Hamilton, 1959) as the main outcome measure (either endpoint score or change from baseline, as presented by the authors). One exception was the study reported by Walkup *et al*. (2001), which used the Pediatric Anxiety Rating Scale (PARS) (The Research Units on Pediatric Psychopharmacology Anxiety Study Group, 2002). We used the efficacy results from the intent-to-treat dataset including the last observation carried forward, except for the study by Bonne *et al*. (2003), which reported results only from study completers. For trials which used more than one dose arm, the arm with the highest efficacy was entered into the analysis. In trials in which more than one active medication was included, we analyzed each drug as belonging to discrete drug categories. As a result, a study might appear twice in the tables (e.g. the study by Hackett *et al*. [2003] included venlafaxine XR versus placebo and diazepam versus placebo).

An ES analysis was conducted by computing standardized mean differences for endpoint and change scores by subtracting means for drug and placebo and dividing by the pooled standard deviation (Cohen, 1988). For studies that only reported standard errors, these values were converted into standard deviations. We also calculated ES when two active medications were included (i.e. active versus active).

To assess the robustness of our analysis we computed the fail-safe *N*, which evaluates the possibility of publication bias. Fail-safe *N* represents the number of studies with negative findings that would need to be combined with the studies reviewed to lead to a nonsignificant result. The larger the fail-safe *N*, the less likely it is that unpublished studies or future studies would overturn our results (Cooper and Rosenthal, 1990).

Homogeneity of the samples was tested by means of heterogeneity analysis (Q statistic). In the presence of statistically significant heterogeneity, data were inspected for outliers and a fixed-effects model was computed to examine significant differences between studies.

Some studies presented results as endpoint values ($n = 11$), whereas other studies presented the change from baseline ($n = 10$). Baseline values were examined and found to be equivalent for drug and placebo, and we therefore additionally presented all study results combining endpoint and change from baseline ($n = 21$).

Results

Twenty-one clinical trials were identified (Table 1). Five studies of venlafaxine XR (SNRI) were identified (Davidson *et al*., 1999; Rickels *et al*., 2000; Gelenberg *et al*., 2000; Allgulander *et al*., 2001; Hackett *et al*., 2003), of which two included an active comparator arm: buspirone (Davidson *et al*., 1999) and diazepam (Hackett *et al*., 2003). Eight trials studied four SSRIs, all versus placebo (paroxetine $n = 2$ [Pollack *et al*., 2001; Rickels *et al*.,

Table 1 Description of studies included in the analysis

Medication Authors, year	DSM criteria Number of outcome measures (OM)	Number of treatment arms Description of each arm/dose	Length of treatment Flexible vs. fixed dosing	Sample size Adult vs. Children Single vs. Multicenter
Venlafaxine XR Davidson *et al.*, 1999 (18)	DSM-IV 5 OM	4 arms: Venlafaxine XR 75 or 150 mg/day, Buspirone 90 mg/day and Placebo	8 weeks Fixed	*n* = 365 Adults Multicenter
Venlafaxine XR Rickels *et al.*, 2000 (19)	DSM IV 3 OM	4 arms: Venlafaxine XR 75, 150 mg/day or 225 mg/day and Placebo	8 weeks Fixed	*n* = 349 Adults Multicenter
Venlafaxine XR Gelenberg *et al.*, 2000 (20)	DSM IV 4 OM	2 arms: Venlafaxine XR (75–225 mg/day) and Placebo	24 weeks Flexible	*n* = 238 Adults Multicenter
Venlafaxine XR Allgulander *et al.*, 2001 (21)	DSM IV 5 OM	4 arms: Venlafaxine XR 37.5, 75, or 150 mg/day and Placebo	24 weeks Fixed	*n* = 528 Adults Multicenter
Venlafaxine XR Hackett *et al.*, 2003 (15)	DSM IV 3 OM	4 arms: Venlafaxine 75 mg/day or 150 mg/day, Diazepam 15 mg/day or Placebo	8 weeks Fixed	*n* = 539 Adults Multicenter
Paroxetine Pollack *et al.*, 2001 (22)	DSM IV 4 OM	2 arms: Paroxetine (20–50 mg/day) and Placebo	8 weeks Flexible	*n* = 324 Adults Multicenter
Paroxetine Rickels *et al.*, 2003 (23)	DSM IV 4 OM	3 arms: Paroxetine 20 mg/day, Paroxetine 40 mg/day and Placebo.	8 weeks Fixed	*n* = 565 Adults Multicenter
Sertraline Rynn *et al.*, 2001 (8)	DSM-IV 5 OM	2 arms: Sertraline 50 mg/day and Placebo	9 weeks Fixed	*n* = 22 Children/Adolescents Single Center
Sertraline Allgulander *et al.*, 2004 (26)	DSM-IV 7 OM	2 arms: Sertraline 50–150 mg/day and Placebo	12 weeks Flexible	*n* = 370 Adults Multicenter
Escitalopram MD 05 (unpublished data from Forest)	DSM IV 4 OM	2 arms: Escitalopram 10–20 mg/day and Placebo	8 weeks Flexible	*n* = 252 Adults Multicenter
Escitalopram MD 06 (unpublished data from Forest)	DSM IV 4 OM	2 arms: Escitalopram 10–20 mg/day and Placebo	8 weeks Flexible	*n* = 281 Adults Multicenter
Escitalopram Davidson *et al.*, 2004 (24)	DSM IV 4 OM	2 arms: Escitalopram 10–20 mg/day and Placebo	8 weeks Flexible	*n* = 307 Adults Multicenter
Fluvoxamine Walkup *et al.*, 2001 (9)	DSM IV 4 OM	2 arms: Fluvoxamine (up to 300 mg/day) and Placebo	8 weeks Flexible	*n* = 73 Children/adolescents Multicenter
Hydroxyzine Ferreri *et al.*, 1995 (27)	DSM III-R 6 OM	2 arms: Hydroxyzine 50 mg/day and Placebo	4 weeks Fixed	*n* = 110 Adults Multicenter

Hydroxyzine Lader and Scotto, 1998 (28)	DSM IV (75% had mixed depressive and anxious symptoms) 6 OM	3 arms: Hydroxyzine 50 mg/day, Buspirone 20 mg/day and Placebo	4 weeks Fixed	$n = 244$ Adults Multicenter
Hydroxyzine Llorca *et al.*, 2002 (29)	DSM IV 4 OM	3 arms: Hydroxyzine 50 mg/day, Bromazepam 6 mg/day and Placebo	12 weeks Fixed	$n = 210$ Adults Multicenter
Pregabalin Pande *et al.*, 2003 (30)	DSM-IV 5 OM	4 arms: Pregabalin 150 mg/day or 600 mg/day, Lorazepam 6 mg/day and Placebo	4 weeks Fixed	$n = 262$ Adults Multicenter
Pregabalin Feltner *et al.*, 2003 (31)	DSM-IV 5 OM	4 arms: Pregabalin 150 mg/day or 600 mg/day, Lorazepam 6 mg/day and Placebo	4 weeks Fixed	$n = 271$ Adults Multicenter
Buspirone Lader and Scotto, 1998 (28)	DSM IV (75% had mixed depressive and anxious symptoms) 6 OM	3 arms: Hydroxyzine 50 mg/day, Buspirone 20 mg/day and Placebo	4 weeks Fixed	$n = 244$ Adults Multicenter
Buspirone Davidson *et al.*, 1999 (18)	DSM-IV 5 OM	4 arms: Venlafaxine XR 75 or 150 mg/day, Buspirone 90 mg/day and Placebo	8 weeks Fixed	$n = 365$ Adults Multicenter
Alprazolam Moller *et al.*, 2001 (33)	ICD-10 5 OM	3 arms: Opipramol 200 mg/day, Alpraaolam 2 mg/day and Placebo	4 weeks Fixed	$n = 307$ Adults Multicenter
Lorazepam Pande *et al.*, 2003 (30)	DSM-IV 5 OM	4 arms: Pregabalin 150 mg/day or 600 mg/day, Lorazepam 6 mg/day and Placebo	4 weeks Fixed	$n = 262$ Adults Multicenter
Lorazepam Feltner *et al.*, 2003 (31)	DSM-IV 5 OM	4 arms: Pregabalin 150 mg/day or 600 mg/day, Lorazepam 6 mg/day and Placebo	4 weeks Fixed	$n = 271$ Adults Multicenter
Diazepam Hackett *et al.*, 2003 (15)	DSM IV 3 OM	4 arms: Venlafaxine 75 mg/day or 150 mg/day, Diazepam 15 mg/day or Placebo	8 weeks Fixed	$n = 539$ Adults Multicenter
Homeopathy Bonne *et al.*, 2003 (14)	DSM IV 7 OM	2 arms: Homeopathic preparations and Placebo	10 weeks Flexible	$n = 39$ Adults Multicenter
Kava-Kava Connor and Davidson, 2002 (32)	DSM IV 4 OM	2 arms: Kava-Kava 280 mg kl/day and Placebo	4 weeks Fixed	$n = 35$ Adults Single center
Opipramol Moller *et al.*, 2001 (33)	ICD-10 5 OM	3 arms: Opipramol 200 mg/day, Alpraaolam 2 mg/day and Placebo	8 weeks Fixed	$n = 365$ Adults Multicenter
OVERALL 21 studies	DSM III-R: 1study; DSM IV: 19 studies; ICD-10: 1 study OM: 3–7	2–4 arms 7 studies included an active comparator	4–24 weeks Fixed/Flexible: 14/7 studies	Total $n = 5935$ Adults/Children: 19/2 studies Multicenter/Single center: 19/2 studies

2003]; escitalopram $n = 3$ [Davidson *et al*., 2004; Goodman *et al*., 2005 and unpublished data from Forest Laboratories]; sertraline $n = 2$ [Rynn *et al*., 2001; Allgulander *et al*., 2004]; and fluvoxamine $n = 1$ [Walkup *et al*., 2001]). Of three AH studies (Ferreri *et al*., 1995; Lader and Scotto, 1998; Llorca *et al*., 2002), two included active comparators: buspirone (Lader and Scotto, 1998) and bromazepam (Llorca *et al*., 2002), although the efficacy data for bromazepam was not provided and, hence, not included in the analysis. Two studies evaluated the anticonvulsant PBG, with lorazepam as active comparator (Pande *et al*., 2003; Feltner *et al*., 2003). Two CAM studies were identified, using kava-kava and homeopathy (Connor and Davidson, 2002; Bonne *et al*., 2003). In one study of alprazolam and opipramol (a tricyclic iminostilbene derivative) (Möller *et al*., 2001), we did not include the tricyclic, as it was the only tricyclic identified in our series.

In summary, we included eight SSRI studies, five venlafaxine studies (SNRI), three hydroxyzine studies (AH), two PGB studies, two buspirone arms, four BZ arms and two CAM studies (kava-kava and homeopathy). Nineteen trials assessed adult patients and two assessed children or adolescents (sertraline by Rynn *et al*. [2001] and fluvoxamine by Walkup *et al*. [2001]). In most trials, GAD was the primary diagnosis, except for the study by Walkup and colleagues, which included children with other anxiety disorders (social phobia and separation anxiety disorder). However, in the latter study we analysed the data from the 73 subjects with GAD, as provided by the authors (B. Vitiello, M.D. personal communication, November, 2003). See Table 1 for a description of these studies.

Baseline values

Examination of baseline group assignment for all of the studies ($n = 20$ using the HAM-A and 1 study using the PARS) yielded a mean ES $\pm$ 0.95 SE confidence limits of 0.04 ± 0.06, which was not significant ($p > 0.19$) and indicates that no bias in group assignment existed prior to the evaluation of drug effects. Therefore, ensuing analyses of endpoint measures were not confounded by baseline differences between groups.

Baseline and change/endpoint values are provided in Table 2.

Sample homogeneity and effect sizes

Combining all studies which reported change in score and endpoint scores yielded an ES of 0.39 ± 0.06, which was highly significant ($p < 0.0001$). The Q statistic was 51.77, which was significant ($p < 0.01$), indicating that the distribution did not estimate a common population mean. Inspection of the ESs revealed four outliers: two negative ESs yielded by Bonne *et al*. (2003) (-0.09) and Connor and Davidson (2002) (-0.88) and two strongly positive ESs yielded by Rynn *et al*. (2001) (1.86) and Walkup *et al*. (2001) (1.26) in contrast to the remaining studies, which ranged between 0.14 and 0.61. Notably, both of the studies yielding negative effect sizes were studies of CAM, whereas both of the studies yielding extremely large ESs were in children and adolescents. Excluding these outliers resulted in an ES of 0.38 ± 0.06 ($p < 0.0001$) but reduced the Q statistic to 22.05, which was no longer significant.

ES for each drug category

As shown in Table 3, for each drug category the mean ESs from highest to lowest were: 1) PGB, 2) hydroxyzine, 3) venlafaxine XR, 4) BZs, 5) SSRI, 6) buspirone and 7) CAM. All ESs were highly significant relative to placebo with the exception of buspirone and the CAM compounds, which did not differ.

We also evaluated the effect size of differences between pairs of active drugs in studies which included an active comparator. The effect sizes were small, as follows: venlafaxine XR versus buspirone (Davidson *et al*, 1999), 0.20; venlafaxine XR versus diazepam (Hackett *et al*., 2003), 0.07; hydroxyzine versus buspirone (Lader and Scotto, 1998), 0.26; pregabalin versus lorazepam (Pande *et al*., 2003), -0.16; pregabalin versus lorazepam (Feltner *et al*., 2003), 0.22; opipramol versus alprazolam (Möller *et al*., 2001), -0.07.

Subgroup analysis

Pharmaceutical versus CAM Examination of pharmaceutical compounds versus CAM for studies examining HAM-A endpoint scores revealed a Q statistic between groups of 8.50, which was significant ($p < 0.01$), indicating that the two treatments were not equivalent. Combining all change and endpoint scores yielded nonsignificant Q values within each group (pharmaceutical = 22.05, homeopathy = 1.14) and a Q between groups of 8.47, which was significant ($p < 0.01$).

The combined change/endpoint ES for the two CAM studies was -0.31 ± 0.46, consistent with drug performing worse than placebo.

Adults versus children The Q statistic for studies examining HAM-A endpoint scores between groups was 17.88, which was highly significant ($p < 0.001$), indicating that the two types of subject populations responded differently to drug. Combining all change and endpoint scores (excluding homeopathy studies) yielded nonsignificant Q values (adults = 22.05, children = 1.10), indicating homogeneity within each group. The Q between groups was 18.61, which was highly significant ($p < 0.001$).

The combined change-from-baseline-and-endpoint ES for the two studies of children and adolescents was 1.38 ± 0.45, which was highly significant ($p < 0.0001$).

Effect of other variables on ES

Examination of the effect of different variables on the ES (i.e. number of study arms, fixed versus flexible dosing, date of publication, location of the study and number of outcome measures used) failed to reveal any significant differences.

Publication bias

A fail-safe *N* revealed that approximately 23 unpublished or new negative studies would be needed to reverse our results. A funnel plot is presented in Figure 1.

Table 2 Efficacy results from baseline to endpoint

Outcome measure	Placebo	Active medication	Active comparator	Medication/s dosage/s	Authors, year
HAM-A baseline (SD)	23.7 (4.2)	23.7 (4.1)	23.8 (4.6)	Venlafaxine XR 75 mg/day; Buspirone 30 mg/day (AC)	Davidson *et al.*, 1999 (18)
HAM-A endpoint (SD)	15.41 (7.25)	12.87 (7.55)	14.38 (7.28)		
HAM-A baseline (SD)	24.1 (4.2)	23.6 (3.7)	N/A	Venlafaxine XR 225 mg/day	Rickels *et al.*, 2000 (19)
HAM-A endpoint (SD)	14.60 (8.34)	12.11 (7.46)			
HAM-A baseline (SD)	25.0 (5)	25.0 (5)	NA	Venlafaxine XR 75–225 mg/day	Gelenberg *et al.*, 2000 (20)
HAM-A endpoint (SD)	16.11 (7.14)	12.05 (7.03)			
HAM-A baseline (SD)	26.7 (5.54)	26.3 (4.06)	NA	Venlafaxine XR 150 mg/day	Allgulander *et al.*, 2001 (21)
HAM-A endpoint (SD)	16.51 (8.44)	11.54 (8.36)			
HAM-A baseline (SD)	27.6 (5.34)	29.97 (5.25)	28.44 (5.16)	Venlafaxine XR 75 mg/day; Diazepam 15 mg/day (AC)	Hackett *et al.*, 2003 (15)
HAM-A endpoint (SD)	15.50 (10.01)	12.99 (9.07)	13.64 (8.57)		
HAM-A baseline (SE)	23.6 (0.30)	23.9 (0.30)	NA	Paroxetine 20–50 mg/day	Pollack *et al.*, 2001 (22)
HAM-A change (SE)	−9.5 (0.7)	−11.8 (0.7)			
HAM-A baseline (SD)	24.4 (3.7)	24.1 (3.6)	NA	Paroxetine 20 mg/day	Rickels *et al.*, 2003 (23)
HAM-A change (SD)	−9.3 (8.7)	−12.5 (8.4)			
HAM-A baseline (SD)	23.3 (4.0)	20.6 (3.6)	NA	Sertraline 50 mg/day	Rynn *et al.*, 2001 (8)
HAM-A endpoint (SD)	21.0 (7.8)	7.8 (5.7)			
HAM-A baseline (SE)	24.71 (0.32)	24.27 (0.32)	NA	Sertraline 50–150 mg/day	Unpublished data from Pfizer
HAM-A change (SE)	−7.96 (0.61)	−11.66 (0.62)			
HAM-A baseline (SD)	22.1 (3.7)	22.8 (3.6)	NA	Escitalopram 10–20 mg/day	MD 05 (Data from Forest labs) 2001
HAM-A change (SD)	−7.7 (6.3)	−9.6 (7.1)			
HAM-A baseline (SD)	22.6 (3.5)	22.6 (3.4)	NA	Escitalopram 10–20 mg/day	MD 06 (Data from Forest labs) 2001
HAM-A change (SD)	−7.6 (6.0)	−9.2 (6.5)			
HAM-A baseline (SD)	23.2 (4.2)	23.6 (4.6)	NA	Escitalopram 10–20 mg/day	Davidson *et al.*, 2004 (24)
HAM-A change (SD)	−7.4 (7.1)	−11.3 (7.3)			
PARS baseline (SD)	19.39 (3.10)	18.74 (3.40)	NA	Fluvoxamine up to 300 mg/day	Walkup *et al.*, 2001 (9)
PARS endpoint (SD)	17.07 (5.14)	9.47 (6.89)			
HAM-A baseline (SD)	24.1 (3.9)	25.9 (4.2)	NA	Hydroxyzine 50 mg/day	Ferreri *et al.*, 1999 (27)
HAM-A change (SD)	−6.6 (7.4)	−11.5 (7.1)			
HAM-A baseline (SD)	26.2 (4.2)	26.6 (4.3)	26.7 (4.1)	Hydroxyzine 50 mg/day; Buspirone 20 mg/day (AC)	Lader and Scotto, 1998 (28)
HAM-A change (SD)	−7.2 (7.7)	−10.8 (7.5)	−8.8 (7.8)		
HAM-A baseline (SD)	25.73 (4.14)	25.49 (3.61)	25.32 (3.44)	Hydroxyzine 50 mg/day; Bromazepam 6 mg/day (AC)	Llorca *et al.*, 2002 (29)
HAM-A change (SD)	−9.64 (7.74)	−12.16 (7.74)	Not available		
HAM-A baseline (SD)	22.9 (3.88)	23.16 (2.73)	23.85 (3.24)	Pregabalin 600 mg/day; Lorazepam 6 mg/day (AC)	Pande *et al.*, 2003 (30)
HAM-A endpoint (SD)	16.23 (7.00)	13.04 (6.89)	11.92 (7.16)		
HAM-A baseline (SD)	24.8 (4.1)	25.4 (4.6)	24.7 (3.7)	Pregabalin 600 mg/day; Lorazepam 6 mg/day (AC)	Feltner *et al.*, 2003 (31)
HAM-A change (SE)	−9.27 (0.84)	−13.17 (0.88)	−11.62 (0.85)		
HAM-A baseline (SD)	29.3 (7.0)	27.7 (7.4)	27.9 (7.6)	Opipramol 200 mg/day; Alprazolam 2 mg/day (AC)	Moller *et al.*, 2001 (33)
HAM-A endpoint (SD)	16.2 (8.7)	13.1 (7.7)	12.6 (7.5)		
HAM-A baseline (SD)	30.4 (7.6)	31.4 (7.2)	NA	Homeopathic preparations	Bonne *et al.*, 2003 (14)
HAM-A endpoint (SD)	20.9 (9.2)	21.7 (11.6)			
HAM-A baseline (SD)	18.8 (2.9)	19.9 (4.1)	NA	Kava-Kava 280 mg kl/day	Connor and Davidson, 2002 (32)
HAM-A endpoint (SD)	10.3 (4.4)	14.2 (8.3)			

HAM-A: Hamilton Anxiety Scale; PARS: Pediatric Anxiety Rating Scale; SD: standard deviation; SE: standard error; NA: not applicable; AC: active comparator.

Table 3 Effect sizes by medication

Medication	ES ± SD	*p*-Value
Pregabalin	0.50 ± 0.24	$p < 0.0001$
Hydroxyzine	0.45 ± 0.18	$p < 0.0001$
Venlafaxine XR	0.42 ± 0.12	$p < 0.0001$
Benzodiazepines	0.38 ± 0.15	$p < 0.0001$
SSRI	0.36 ± 0.09	$p < 0.0001$
Buspirone	0.17 ± 0.21	NS
CAM	−0.31 ± 0.46	NS
All	0.39 ± 0.06	$p < 0.0001$
All (with outliers removed)	0.38 ± 0.06	$p < 0.0001$

ES: effect size; SD: standard deviation; NS: non-significant.

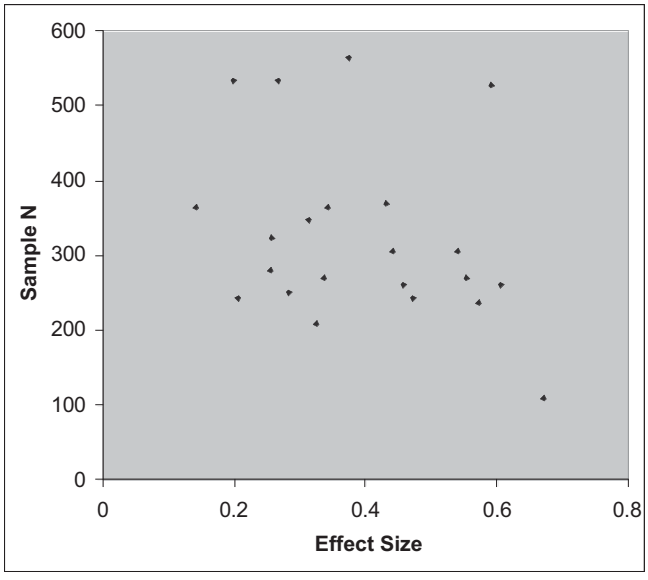

Figure 1 Funnel plot of endpoint effect size as a function of sample size.

Discussion

Our analysis showed a low to moderate overall ES (0.39 ± 0.06) for drug therapy in the treatment of GAD. Only type of treatment (CAM versus conventional) and age of subjects (children/adolescents versus adults) influenced the ES. The insignificant effects of study design (number of arms, fixed versus flexible dose, date or location of trial) provide some confidence that the effects of real treatment vs. placebo are less likely to be obscured by these confounders, unlike some depression trials (Khan *et al*., 2000, 2003; Zimmerman and Posternak, 2003; Walsh *et al*., 2002).

Although moderate, this mean ES is, nevertheless, slightly higher compared with the ES reported by the NICE group for treatment of major depression with SSRIs: 0.34 (NICE, 2004). When the active drugs were compared, all effect sizes were small and no consistent pattern was seen.

Compared with other anxiety disorders, GAD had a lower response to treatment in terms of ES: obsessive-compulsive disorder, 0.45–1.48 (Greist *et al*., 1995); panic disorder, 0.55 (Otto *et al*., 2001); social phobia, 0.14–1.12 (Hidalgo *et al*., 2001) and was somewhat similar to PTSD, 0.42 for paroxetine and 0.26 for sertraline (NICE, 2005).

Our ESs are also slightly higher that the ones reported by Mitte *et al*. (2005) in their meta-analytic review of the drug treatment of GAD (mean random ES for HAM-A g = 0.33, 95% CI 0.27–0.39). However, Mitte *et al*.'s work had several methodological differences from our present investigation. For example, they used broader diagnostic criteria, including earlier DSM versions (DSM-II, DSM-III), and they included studies evaluating treatment of comorbid anxiety and depression and reports for which we did not have enough information in the publication to permit calculation. In addition, Mitte *et al*. focused their analysis primarily on AZAs and BZs, including only few AHs and SSRIs, and they did not include trials on children/adolescent populations or CAM.

In our analysis PGB and hydroxyzine demonstrated the highest ES followed by venlafaxine XR, BZs and SSRIs. On the other hand, buspirone was not significantly different from placebo and had a low ES.

Interpretation of the higher ES for PGB and hydroxyzine is not straightforward, but a number of issues can be considered. First, the PGB and AH studies tended to be short, but whether these drugs have faster onset of action relative to antidepressants is unknown, and we cannot address this question due to the lack of short-term antidepressant studies in our sample. Second, an important clinical question is whether the effects of PGB and hydroxyzine would remain the same after more weeks or months of treatment. Long-term (i.e. 8–12 weeks) PGB and AH studies are needed.

Methodological differences between the studies included may have influenced the magnitude of the ES found for the different treatments. Some of these differences may not have been clearly expounded in the authors' reports (i.e. rater training, day-to-day application of the inclusion/exclusion criteria, and so on).

Another interesting finding was the strikingly different ESs in two pairs of groups: 1) adults versus children and adolescents and 2) conventional pharmacotherapy versus CAM. The two studies conducted in children and adolescents yielded larger ESs than adult studies. This difference between adult and child/adolescent ESs was statistically significant. However, it is difficult to interpret the clinical meaning of this finding. The relatively small samples in these two studies are a limiting factor. On the other hand, such a strong response to SSRIs in children and adolescents with GAD may suggest that these compounds were clearly effective in treating childhood/adolescent anxiety in these two studies. We know that GAD is chronic and that when adults present for treatment of GAD they often have been suffering from the disorder for an average of 10–15 years (Eisen, 1998; Yonkers *et al*., 1996). Therefore, it is tempting to consider that this high ES in children and adolescents indicates that the sooner the disorder

receives treatment, the higher the likelihood of a marked response to therapy.

Regarding the comparison of conventional treatment versus CAM, interestingly, CAM therapy had a negative mean ES (-0.31 ± 0.46) compared with placebo. In other words, patients experiencing GAD taking a sugar pill did better than patients treated with kava-kava or homeopathy. In our view, this is an important finding with potential implications for the public-health arena, considering that these compounds are readily available and loosely regulated by the FDA. Moreover, kava-kava is now associated with serious liver toxicity (Stickel *et al*., 2003), yet it still remains available in the US market. Furthermore, the two studies were conducted in different countries yet still produced discouraging results for these particular compounds. Nevertheless, we must be conservative in our interpretation of this finding, as the pooled sample treated with CAM consisted of only 74 subjects and used only two single compounds (kava-kava and homeopathic preparation). Variables that did not have a significant effect on outcome were: number of study arms, fixed versus flexible dosing, length of treatment, location of the site and number of outcome measures used.

We do not believe that our results would have been materially different if unpublished negative studies of these treatments were in existence. Twenty-three such studies would be needed to overturn our findings.

Conclusions

Medications investigated in double-blind randomized clinical trials included in our meta-analysis varied in the magnitude of their ES, ranging from moderate to poor. On the higher end of the spectrum we found the anticonvulsant PGB and the AH hydroxyzine, followed in order by venlafaxine XR, BZs and SSRIs. Buspirone had a low ES, and CAM (kava-kava and homeopathy) performed worse than placebo. Treatment effect for children and adolescents was greater than for adults.

Conflict of interest statement

Dr Jonathan R. T. Davidson: Speaker for Solvay Pharmaceuticals, Pfizer Inc., GlaxoSmithKline, Wyeth Pharmaceuticals, Lichtwer Pharma, Forest, American Psychiatric Association. Research Support from Pfizer, Solvay, Eli Lilly, GlaxoSmithKline, Wyeth, Organon Inc., Forest, PureWorld, Allergan, Nutrition 21, Bristol Myers Squibb, Johnson and Johnson, Cephalon, AstraZeneca, Parke Davis, Pharmacia, Upjohn, UCB, Merk, and Janssen. Advisory Board for Solvay, Pfizer, GlaxoSmithKline, Forest, Eli Lilly, Ancile, Roche, MediciNova, Jazz, Novartis, Organon, Boehringer Ingelheim, MedTap, Research Triangle Institute, AstraZeneca, Johnson and Johnson, Wyeth, Bristol Myers Squibb, Boots, UCB, Sanofi-Synthelabo, Alexza, and Janssen. Drugs Supplied for Other Studies from Eli Lilly, Schwabe, PureWorld and Pfizer Inc. Royalties from Multi-Health Systems Inc., Guildford Publications, American Psychiatric Association, Penguin Putnam Publishers, Current Medical Science, Martin Dunitz, and Taylor and Francis.

Dr Rosario B. Hidalgo and Dr Larry A. Tupler do not have any conflict of interest to disclose.

References

Allgulander C, Hackett D, Salinas E (2001) Venlafaxine extended release (ER) in the treatment of generalised anxiety disorder: twenty-four-week placebo-controlled dose-ranging study. Br J Psychiatry *179*: 15–22

Allgulander C, Dahl A A, Austin C, Morris P L P, Sogaard J A, Fayyad R, Kutcher S P, Clary C M (2004) Efficacy of sertraline in a 12-week trial for generalized anxiety disorder Am J Psychiatry *161*: 1642–1649

Bonne O, Shemer Y, Gorali Y, Katz M, Shalev A Y (2003) A randomized, double-blind, placebo-controlled study of classical homeopathy in generalized anxiety disorder. J Clin Psychiatry *64*: 282–287

Cohen J (1988) Statistical Power Analysis for the Behavioral Sciences (2nd edition). Lawrence Earlbaum Associates, Hillsdale, NJ, USA

Connor K M and Davidson J R T (2002) A placebo-controlled study of kava kava in generalized anxiety disorder. Int Clin Psychopharmacology *17*: 186–188

Cooper H M, Rosenthal R (1990) Statistical versus traditional procedures for summarizing research findings. Psychol Bull *87*: 442–449

Davidson J R T, DuPont R L, Hedges D, Haskins J T (1999) Efficacy, safety and tolerability of venlafaxine extended release and buspirone in outpatients with generalized anxiety disorder. J Clin Psychiatry *60*: 528–535

Davidson J R T, Bose A, Korotzer A, Zheng H (2004) Escitalopram in the treatment of generalized anxiety disorder: double-blind, placebo controlled, flexible-dose study. Depression and Anxiety *19*: 234–240

Eisen J L (1998) Course and psychosocial impairment in GAD. Presented at *151st Annual Meeting of the American Psychiatric Association*

Feltner D E, Crockatt J G, Dubovsky S J (2003) A randomized, double-blind, placebo-controlled, fixed dose, multicenter study of pregabalin in patients with generalized anxiety disorder. J Clin Psychopharmacol *23*: 240–249

Ferreri M, Darcis T, Burtin B, and the French Study Group for Hydroxyzine (1995) A multicentre double-blind placebo-controlled study investigating the anxiolytic efficacy of hydroxyzine in patients with generalized anxiety. Human Psychopharmacology *10*: 181–187

Gelenberg A, Lydiard B R, Rudolph R, Aguiar L, Haskins J T, Salinas E (2000) Efficacy of venlafaxine extended release capsules in nondepressed outpatients with generalized anxiety disorder: a 6-month randomized controlled trial. JAMA *283*: 3082–3088

Goodman W K, Bose A, Wang Q (2005) Treatment of generalized anxiety disorder with escitalopram: pooled results from double-blind, placebo-controlled trials. J Affect Disord *87*: 161–167

Greist J H, Jefferson J W, Kobak K A, Katzelnick D J, Serlin R C (1995) Efficacy and tolerability of serotonin transport inhibitors in obsessive-compulsive disorder: a meta-analysis. Arch Gen Psychiatry *52*: 53–60

Hackett D, Haudiquet V, Salinas E (2003) A method for controlling for a high placebo response rate in a comparison of venlafaxine XR and diazepam in the short-term treatment of patients with generalised anxiety disorder. Eur Psychiatry *18*: 182–187

Hamilton M (1959) The assessment of anxiety states by rating. Br J Med Psychology *32*: 50–55

Hidalgo R B, Barnett S D, Davidson J R T (2001) Social anxiety disorder in review: two decades of progress. International J of Neuropsychopharmacol *4*: 279–298

Hidalgo R B, Davidson J R T (2001) Generalized anxiety disorder: an important clinical concern. Med Clin of North Am *85*: 691–710

Khan A, Warner H, Brown W A (2000) Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials: an

analysis of the Food and Drug Administration database. Arch Gen Psychiatry *57*: 311–317

Khan A, Khan S R, Walens G, Kolts R, Giller E (2003) Frequency of positive studies among fixed and flexible dose antidepressant clinical trials: an analysis of the Food and Drug Administration summary basis of approval protocols. Neuropsychopharmacology *28*: 552–557

Lader M, Scotto J C (1998) A multicentre double-blind comparison of hydroxyzine, buspirone and placebo in patients with generalized anxiety disorder. Psychopharmacology *139*: 402–406

Llorca P M, Spadone C, Sol O, Danniau A, Bougerol T, Corruble E, Faruch M, Macher J P, Sermet E, Servant D (2002) Efficacy and safety of hydroxyzine in the treatment of generalized anxiety disorder: a 3-month double-blind study. J Clin Psychiatry *63*: 1020–1027

Mitte K, Noack P, Steil R, Hautzinger M (2005) A meta-analytic review of the efficacy of drug treatment in generalized anxiety disorder. J Clin Psychopharmacology *25*: 141–150

Möller H-J, Volz H-P, Reimann I W, Stoll K-D (2001) Opipramol for the treatment of generalized anxiety disorder: a placebo-controlled trial including an alprazolam-treated group. J Clin Psychopharmacol *21*: 59–65

National Institute for Health and Clinical Excellence (NICE) (December 2004) Clinical guideline on depression. CG23 Depression: Appendix 20. Available online at: http://www.nice.org.uk/. Accessed 18 March 2006

National Institute for Health and Clinical Excellence (NICE) (March 2005) Clinical guideline on PTSD. CG26 PTSD: Appendix 16b. Available online at: http://www.nice.org.uk/. Accessed 18 March 2006

Otto M W, Tuby K S, Gould R A, McLean R Y S, Pollack M H (2001) An effect-size analysis of the relative efficacy and tolerability of serotonin selective reuptake inhibitors for panic disorder. Am J Psychiatry *158*: 1989–1992

Pande A C, Crockatt J G, Feltner D E, Janney C A, Smith W T, Weisler R, Londborg P D, Bielski R J, Zimbroff D L, Davidson J R T, Liu-Dumaw M (2003) Pregabalin in generalized anxiety disorder: a placebo-controlled trial. Am J Psychiatry *160*: 533–540

Pittler M H, Ernst E (2000) Efficacy of kava extract for treating anxiety: systematic review and meta-analysis. J Clin Psychopharmacol *20*: 84–89

Pollack M H, Zaninelli R, Goddard A, McCafferty J P, Bellew K M, Burnham D B, Iyengar M K (2001) Paroxetine in the treatment of generalized anxiety disorder: results of a placebo-controlled, flexible-dosage trial. J Clin Psychiatry *62*: 350–357

Rickels K, Pollack M H, Sheehan D V, Haskins J T (2000) Efficacy of extended-release venlafaxine in nondepressed outpatients with generalized anxiety disorder. Am J Psychiatry *157*: 968–974

Rickels K, Rynn M A (2001) What is generalized anxiety disorder? J Clin Psychiatry *62*: 4–12

Rickels K, Zaninelli R, McCafferty J, Bellew K, Iyengar M, Sheehan D (2003) Paroxetine treatment of generalized anxiety disorder: a double-blind, placebo controlled study. Am J Psychiatry *160*: 749–756

Rynn M A, Siqueland L, Rickels K (2001) Placebo-controlled trial of sertraline in the treatment of children with generalized anxiety disorder. Am J Psychiatry *158*: 2008–2014

Stickel F, Baumuller H M, Seitz K, Vasilakis D, Seitz G, Seitz H K, Schuppan D (2003) Hepatitis induced by Kava (Piper methysticum rhizona). J Hepatol *39*: 62–66

The Research Units on Pediatric Psychopharmacology Anxiety Study Group (2002) The Pediatric Anxiety Rating Scale (PARS): development and psychometric properties. J of the Am Academy of Child & Adolescent Psychiatry *41*: 1061–1069

Walkup J T, Labellarte M J, Riddle M A, Pine D S, Greenhill L, Klein R, Davies M, Sweeney M, Abikoff H, Hack S, Klee B, McCracken J, Bergman L, Piacentini J, March J, Compton S, Robinson J, O'Hara T, Baker S, Vitiello B, Ritz L A, Roper M (2001) Fluvoxamine for the treatment of anxiety disorders in children and adolescents. N Engl J Med *344*: 1279–1285

Walsh B T, Seidman S N, Sysko R, Gould M (2002) Placebo response in studies of major depression. JAMA *287*: 1840–1847

Wittchen H-U, Zhao S, Kessler R C, Eaton W W (1994) DSM-III-R generalized anxiety disorder in the National Comorbidity Survey. Arch of Gen Psychiatry *51*: 355–364

Yonkers K A, Warshaw M G, Massion A O, Keller M B (1996) Phenomenology and course of generalised anxiety disorder. Br J Psychiatry *168*: 308–313

Zimmerman M, Posternak M A (2003) Placebo response in antidepressant efficacy trials: relationship to number of active treatment groups, in *2003 Annual Meeting New Research Program and Abstracts*. Arlington, VA, American Psychiatric Association, *number 893*