Efficacy and Tolerability of Benzodiazepines versus Antidepressants in Anxiety Disorders: A Systematic Review and Meta-Analysis

Emanuela Offidani a, Jenny Guidi a, Elena Tomba a, Giovanni Andrea Fava a, b

a Affective Disorder Program, Department of Psychology, University of Bologna, Bologna, Italy; b Department of Psychiatry, State University of New York at Buffalo, Buffalo, N.Y., USA

Key Words
Benzodiazepines · Antidepressive agents · Anxiety · Generalized anxiety disorder · Panic · Agoraphobia · Social phobia · Obsessive-compulsive disorder · Meta-analysis · Systematic review

Abstract
Background: Placebo-controlled trials showed that both benzodiazepines (BDZ) and antidepressant drugs (AD) are effective in treating anxiety disorders. However, in the last years a progressive shift in the prescribing pattern from BDZ to newer AD has taken place. The aim of this systematic review and meta-analysis is to analyze whether controlled comparisons support such a shift. Methods: CINHAL, the Cochrane Library, MEDLINE, PubMed and Web of Science were searched from inception up to December 2012. A total of 22 studies met the criteria for inclusion. They were mostly concerned with tricyclic antidepressants (TCA; 18/22) and involved different anxiety disorders. In order to reduce clinical heterogeneity, only the 10 investigations that dealt with the comparison between TCA and BDZ in panic disorder were submitted to meta-analysis, whereas the remaining papers were individually summarized and critically examined. Results: According to the systematic review, no consistent evidence emerged supporting the advantage of using TCA over BDZ in treating generalized anxiety disorder (GAD), complex phobias and mixed anxiety-depressive disorders. Indeed, BDZ showed fewer treatment withdrawals and adverse events than AD. In panic disorder with and without agoraphobia our meta-analysis found BDZ treatments more effective in reducing the number of panic attacks than TCA (risk ratio, RR = 1.13; 95% CI = 1.01–1.27). Furthermore, BDZ medications were significantly better tolerated than TCA drugs, causing less discontinuation (RR = 0.40; 95% CI = 0.20–0.57) and side effects (RR = 0.41; 95% CI = 0.34–0.50). As to newer AD, BDZ trials resulted in comparable or greater improvements and fewer adverse events in patients suffering from GAD or panic disorder. Conclusions: The change in the prescribing pattern favoring newer AD over BDZ in the treatment of anxiety disorders has occurred without supporting evidence. Indeed, the role and usefulness of BDZ need to be reappraised.

Introduction
Anxiety disorders encompassing phobias, panic disorder, generalized anxiety disorder (GAD), obsessive-compulsive disorders and acute and posttraumatic stress disorders are the most common psychiatric conditions, with...
an estimated lifetime prevalence of about 29% [1]. Longitudinal evidence has shown that anxiety disorders usually do not remit over time but rather persist as chronic conditions entailing a substantial economic burden of approximately USD 42 billion in both the USA and Europe [2, 3]. Given the impact of anxiety disorders on mental health and their social and economic costs, a number of placebo-controlled studies were conducted suggesting the efficacy of both benzodiazepines (BDZ) and antidepressant drugs (AD) in treating such conditions [4].

In the last years, a progressive change in the prescribing pattern from BDZ to newer antidepressants (SSRI, SNRI) has been observed [5–7]. In 2008, Berney et al. [8] published a systematic review of controlled comparisons between BDZ and AD trials in anxiety disorders up to 2003. They were able to identify only 1 trial comparing diazepam with the new AD, venlafaxine XR [9], concluding that such a shift in drug treatment of anxiety disorders occurred without any comparative evidence [8]. The aim of this paper was to update the systematic review by Berney et al. [8] on controlled direct comparisons between AD and BDZ in anxiety disorders, applying quantitative methods when feasible.

### Methods

#### Data Sources

PRISMA guidelines were used to conduct the systematic review of the literature for identifying randomized controlled trials comparing BDZ and AD in the treatment of anxiety disorders [10]. Key words were 'anxiety', 'benzodiazepine' and 'antidepressant'. To increase identification of studies involving anxiety disorders, we expanded the search terms to include 'obsessive-compulsive', 'OCD', 'generalized anxiety', 'GAD', 'phobia' and 'social phobia'. Limits were set to randomized controlled trials and adult trials in the English language. Electronic research-literature databases included CINAHL, the Cochrane Library, PubMed and Web of Science from the inception of each database to December 2012. In addition, reference lists of initially identified reports were examined and further clinical trials were searched manually.

#### Study Selection

Searching and ratings of target responses were carried out independently by two investigators (E.O. and J.G.); disagreements were resolved by consensus among these primary raters and a senior investigator (E.T. or G.A.F.). We selected (1) randomized controlled trials examining (2) the efficacy of treatment with BDZ versus AD of (3) adult patients (4) with anxiety disorders. The primary outcome measures were response rates as defined by the study investigators (i.e. being free of panic attacks or being judged as improved by the clinician). Secondary outcome measures were dropout rates and occurrence of adverse events (AE) during treatments.

We excluded studies if they: (1) were not randomized controlled trials, (2) focused on treatment of patients with primary diagnoses other than anxiety disorders or (3) were conducted in nonclinical samples. We also excluded studies that: (d) involved patients younger than age 18, (5) did not compare directly treatment with BDZ versus AD, (6) did not contain original data or (7) reported outcomes other than treatment efficacy and/or AE during treatment, as well as (8) studies in which response rates were not identified categorically (studies submitted to meta-analysis only).

#### Data Extraction

Data were independently extracted by both reviewers with the use of a precoded form. The following data were extracted from studies meeting criteria for inclusion in the systematic review: (1) age, gender distribution, methods used to define and diagnose study participants and other inclusion criteria at baseline, (2) group comparisons, type of pharmacological treatment, number of patients randomized to each treatment arm, treatment duration and assessment times and (3) methods used to define response to treatment and specification of reasons for dropout and of AE. The methodological quality of the included trials was assessed independently by both reviewers based on three basic criteria: random allocation of treatments, blinding of outcome assessment and handling of attrition.

#### Data Synthesis

When it was judged to be feasible, data were submitted to meta-analysis. The primary outcome of the meta-analysis was treatment efficacy, expressed as response rates. Therefore, the risk ratio (RR) of response and its standard error (SE) were calculated from each study. Examination of the pooled results was performed based on the random-effects model to increase the generalizability of findings, since this model is more conservative than the fixed-effects model. An alpha level of 0.05 was used for hypothesis tests.

In addition to point estimates and confidence intervals, the Q statistic was performed to assess heterogeneity between study results. With this statistic the null hypothesis was tested so that effect sizes from each of the studies were similar enough that a common population effect size could be calculated [11]. However, the Q statistic only informs about the presence versus the absence of heterogeneity, but it does not report on the extent of such heterogeneity. The I² statistic, which is an indicator of heterogeneity in percentages, was also calculated. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity, with 25% as low, 50% as moderate and 75% as high heterogeneity [12].

The likelihood of significant publication bias was assessed through Begg’s funnel plot [13] and testing for asymmetry using Egger’s test statistic [14]. Sensitivity analyses were implemented in order to estimate the influence of each study by deleting each in turn from the analysis and noting the degree to which the size and significance of the treatment effect changed. Metaggerson was performed to investigate how certain characteristics (i.e. treatment duration, publication year, presence of a comorbid mood disorder and intention-to-treat, ITT) acted to influence treatment effects. All analyses were conducted using the user-written packages for meta-analysis available in Stata 10.1 (Stata Corporation, College Station, Tex., USA).
Results

Characteristics of Included Studies

The initial search identified 222 reports involving BDZ and AD in the treatment of anxiety disorders for potential inclusion in the systematic review (online suppl. fig. 1; for all online suppl. material, see www.karger.com/doi10.1159/000353198). Of these, we excluded 111 studies, which focused on the treatment of patients with primary diagnoses other than anxiety disorders or were laboratory trials or studies conducted in nonclinical populations. We excluded a further 89 studies because they did not compare directly the treatment of anxiety disorders with BDZ versus AD, represented reanalyses of data published elsewhere, or reported outcomes other than treatment efficacy or adverse effects during treatment.

Therefore, a total of 22 papers met the criteria for inclusion in the study. There were 18 studies concerned with tricyclic antidepressants (TCA), 1 with phenelzine and 3 studies with newer AD; 9 studies compared TCA with BDZ in mixed anxiety, GAD and specific or complex phobias. These studies and those involving newer AD will be summarized and critically reviewed according to the characteristics of each sample.

Only in panic disorder was there a sufficient number of trials to warrant quantitative methods of analysis. We thus submitted to meta-analysis data from 10 reports (11 comparisons) on treatment of panic disorder with or without agoraphobia with BDZ versus TCA.

BDZ versus TCA

Mixed Anxiety

A total of 5 reports involved participants suffering from a broad range of anxiety disturbances (anxiety neurosis, mixed anxiety disorders and mixed anxiety and depressive disorders); 2 studies compared diazepam and clomipramine [15] or dothiepin [16], 2 compared alprazolam and imipramine [17] or amitriptyline [18], and 1 chlordiazepoxide and imipramine [19] (online suppl. table 1).

As to response to treatment, 2 of the 5 studies [15, 19] found that TCA were more effective than BDZ, 1 study [17] reported a significantly better improvement with alprazolam than imipramine, and 1 showed no difference between the two drug classes [18] (online suppl. table 1). Tyrer et al. [16] reported that, according to CPRS (Comprehensive Psychopathological Rating Scale) [20] and MADRAS (Montgomery-Asberg Depression Rating Scale) [21] scores, diazepam resulted in being less effective than other treatments (i.e. dothiepin, cognitive-behavioral therapy, self-help and placebo). However, these findings are difficult to interpret not only because the sample included patients with dysthymia, but also because a percentage between 7 and 31% of participants received additional pharmacological treatments [16].

Dropout percentages were greater among patients treated with TCA than in those who received BDZ. However, only Kahn et al. [19] reported such difference to be significant (online suppl. table 1). As to adverse reactions, only 2 studies reported rates of side effects experienced by patients [15, 18]. Draper and Daly [18] found no differences in rates of AE between alprazolam and amitriptyline, whereas Allsopp et al. [15] suggested a better tolerability of BDZ compared to TCA, with only 17% of subjects taking diazepam (10–30 mg/day) complaining of adverse reactions versus 33% of patients taking clomipramine (25–125 mg/day; online suppl. table 1).

Generalized Anxiety Disorder

A total of 3 reports directly compared BDZ and TCA in patients suffering from GAD [22–24] (online suppl. table 1); 2 studies compared alprazolam to imipramine and opipramol, respectively [22, 24] and 1 compared diazepam, imipramine and trazadone [23] (online suppl. table 1).

As to efficacy, Hoehn-Saric et al. [22], evaluating response to treatment as improvement on the Hamilton Anxiety scale (HAM-A) [25] from baseline to posttreatment, found alprazolam (0.5–6 mg/day) to be significantly more effective in reducing somatic symptoms of anxiety than imipramine (25–200 mg/day). On the contrary, imipramine yielded greater improvements in psychiatric features of anxiety and depression, such as interpersonal sensitivity, hostility and paranoid ideation, than alprazolam [22]. Rickels et al. [23] found that imipramine led to a greater anxiolytic effect than diazepam in psychological symptoms and a comparable effect in the somatic facets. When only patients with low levels of concomitant depression were considered, treatments produced a comparable effect [23]. Similarly, Möller et al. [24] found no significant differences between BDZ and AD in patients with GAD.

Dropout rates were estimated to be approximately similar in both medication classes. Rates of AE suggested a comparable likelihood of reporting side effects for both opipramol and alprazolam [24] and a slightly greater risk in patients treated with imipramine than in those treated with diazepam or alprazolam [22, 23] (online suppl. table 1).
Social Phobia

Gelenter et al. [26] compared alprazolam, phenelzine, cognitive-behavioral therapy and placebo in 65 patients suffering from social phobia (online suppl. table 1). Overall results showed no differences in self-report questionnaires among groups, while patients treated with phenelzine were rated by physicians as more improved in some measures (i.e. work and social disability) than individuals receiving alprazolam. AE were not assessed or reported.

Meta-Analysis of Studies on Treatment of Panic Disorder

Concerning the treatment of panic disorder, 10 reports contributed data for the meta-analysis (online suppl. fig. 1). Since 1 of them reported findings from 2 independent studies [27], we considered 11 group comparisons. The included studies reported response rates and/or dropout rates and/or rates of AE for a total of 2,624 participants (1,010 patients in the BDZ treatment arm and 962 in the TCA arm). Participants averaged 33.7 (SD = 1.64) years of age and 67.3% (range 58–92%) were female. Selected characteristics of the included studies are presented in online supplementary table 2.

Participants were assigned at random to the conditions in all studies, all of which were double-blind trials, except 1 [27] which was a single-blind (evaluator-blind) trial. ITT analyses were performed in 6 reports [27–33], while in the remaining studies completers’ data only were reported.

Response to Treatment. Clinical improvement was addressed in terms of response rates in 8 of 11 comparisons. Thus, 8 studies [27–34] contributed data for this analysis (online suppl. fig. 2).

The pooled RR for response was 1.134 (95% CI = 1.011–1.271) in the random-effects model, suggesting a relative advantage in response to treatment for the BDZ compared to TCA. However, heterogeneity across trials was statistically significant (Q = 477.490; d.f. = 7; p < 0.001). The I² statistic also indicated high heterogeneity (I² = 98%) among the pooled studies. Both visual inspection of Begg’s funnel plot and Egger’s test (p = 0.163) were not suggestive of the presence of publication bias. A sensitivity analysis was performed to determine the contribution of each study to the overall effect size, and 4 studies [28, 31, 33, 34] seemed to markedly influence the observed RR for response. Removing each of them in turn from the analysis, we did not find a significant advantage of BDZ compared to TCA as to response rates, but heterogeneity across trials was still high.

Performing metaregression analyses, we did not find any significant effect of the above-mentioned characteristics (i.e. treatment mood duration, publication year, presence of a comorbid mood disorder and ITT) on response rates among the included studies.

Dropout Rates. Since all studies reported dropout rates, 11 trials contributed data for this analysis [27–36]. Data showed a significant advantage of the use of BDZ compared to TCA (p < 0.001). The pooled RR for dropouts was 0.404 (95% CI = 0.287–0.569) in the random-effects model. Both Q and I² statistics suggested significant heterogeneity among the pooled studies (Q = 2646.731; d.f. = 10; p < 0.001; I² = 99%). Both Begg’s funnel plot and Egger’s test (p = 0.158) did not indicate the presence of publication bias. Sensitivity analysis did not show any significant influence on the pooled RR for dropouts. Metaregression analyses showed a significant effect of publication year (i.e. before vs. after DSM-IV release; coefficient = 0.383, 95% CI = 0.006–0.760). We also tested for treatment duration, presence of a comorbid mood disorder and ITT analyses, but we did not find significant effects on dropout rates among the pooled studies.

Rates of AE. Only 5 studies [27, 30, 31, 33] reported rates of AE and thus contributed data. Analyses showed that patients randomized to BDZ were significantly less likely to experience adverse effects compared to TCA. Across the trials, the pooled RR for AE was 0.412 (95% CI = 0.340–0.499) in the random-effects model. The Q statistic was significant (Q = 95.542; d.f. = 4; p < 0.001) and the I² statistic indicated high variability in rates of AE among the included studies (I² = 96%). Both visual inspection of Begg’s funnel plot and Egger’s test (p = 0.335) were not suggestive of the presence of publication bias. Sensitivity analyses did not yield any study as influencing the observed RR for AE. Performing metaregression analyses, we did not find any significant effect of the selected variables on rates of AE.

BDZ versus Newer Antidepressants

To date, only 3 studies comparing the efficacy of BDZ medications and newer antidepressants in anxiety disorders have been published [9, 37, 38]. Hackett et al. [9] compared diazepam, venlafaxine XR (150 and 75 mg) and placebo in 540 patients with GAD. Results showed no significant differences in response rates between groups. However, discontinuations due to side effects and AE were more frequent in patients taking venlafaxine XR than in those treated with diazepam. Feltner et al. [37] evaluated the efficacy of a 4-week treatment with lorazepam, paroxetine or placebo in 169 GAD subjects. According to HAM-A [25] scores, both active treatments were effective in reducing anxiety-related psychiatric symp-
Benzodiazepines versus Antidepressants in Anxiety Disorders

Recently, Nardi et al. [38] conducted an open-label, 8-week randomized trial comparing the efficacy and safety of clonazepam and paroxetine in 120 patients with panic disorder with and without agoraphobia. Overall, treatment with clonazepam resulted in significantly fewer panic attacks and greater global improvement than paroxetine. Also, participants treated with clonazepam reported fewer AE than those taking paroxetine (73 vs. 95%; p = 0.001). Furthermore, responders (n = 105) entered a 3-year continued monotherapy with either clonazepam or paroxetine, and clonazepam led to a significantly greater clinical improvements and fewer side effects than paroxetine [39].

**Discussion**

Our systematic review found a paucity of studies providing a controlled direct comparison of AD and BDZ in anxiety disorders. Most of the studies (18/22) were concerned with TCA and only 3 with newer antidepressants [9, 37, 38]. The superiority of AD over BDZ in terms of efficacy and tolerability was not supported by the available evidence.

As to TCA, studies with mixed anxiety were difficult to evaluate because of the heterogeneous features of the samples and the confounding effects of depressive symptoms. However, in mixed anxiety, GAD and social phobia, a superior efficacy of TCA did not clearly emerge, while a better tolerability of BDZ was found. There were many methodological problems that require caution in interpreting these findings. Allsopp et al. [15] reported that, in phobic patients, clomipramine was found to be superior to diazepam in treating situational anxiety. However, 40% of phobic patients taking clomipramine were treated at the maximum dosage (150 mg/day), while only 22% of dzapam participants received 30 mg/day. Authors also reported that results were obtained for 50 patients, but analyses were conducted on only 33 patients without performing the ITT procedure. Similarly, Kahn et al. [19] reported that imipramine led to a significantly better improvement in anxiety disorders than chlordiazepoxide. However, this sample was constituted by patients suffering from either anxiety or depressive disorders, while agoraphobic patients were excluded from the analyses. By contrast, 2 reports found alprazolam to be superior to imipramine in treating patients with GAD and complex phobias [17, 22]. Also in these cases, small sample sizes and the absence of ITT analysis do not allow the drawing of any definite conclusion.

Less efficacy and tolerability of TCA over BDZ were found by the meta-analysis of studies of treatment of panic disorder (with or without agoraphobia). It should be noted, however, that primary diagnoses, definitions of response, dropout rates or reporting of AE varied consistently across trials, depending on inclusion criteria, methods of assessment and severity criteria involved.

It is conceivable, even though yet to be tested, that more comprehensive and sensitive methods of assessment [40, 41] than those traditionally endorsed [21, 21, 25] may disclose differential responsiveness. Some samples were highly heterogeneous [15–19] and this might have increased the likelihood of spurious results [42, 43]. The duration of treatments also varied across the included studies, even though metaregression did not show significant effects on the selected outcomes for the studies submitted to meta-analysis. Furthermore, several of the included studies presented statistical and methodological problems that limit the generalizability of results. For example, in some cases, authors did not take into proper account the presence of possible confounding variables such as sociodemographic characteristics or different rates of depressive comorbidity among the drug arms, or they did not perform the ITT analysis (or failed to mention it) [15, 27, 29, 32, 34–36]. Finally, comparisons involved different active compounds that, even though they belonged to the same medication class, may have distinctive pharmacokinetic characteristics and entail different therapeutic responses [16, 19, 24, 26].

As to the comparisons between newer AD and BDZ, one is impressed by the paucity of studies that were published. Only 3 studies included direct comparisons between BDZ and newer AD [9, 37, 38]. Of these, 2 reports [9, 37] focused mainly on secondary outcomes [9, 37], namely, the placebo response in the former and a new instrument validation in the latter. Both studies failed to report significant differences in response to treatment between newer AD (venlafaxine XR 75 and 150 mg/day and paroxetine 20 mg/day, respectively) and BDZ (diazepam 15 mg/day and lorazepam 1.5 mg/day) in GAD patients. In the 3rd report, a randomized naturalistic study of patients with panic disorder with or without agoraphobia, clonazepam (2 mg/day) was proved to be significantly superior to paroxetine (40 mg/day) in both reducing panic attacks and leading to clinical improvement [38].

Our systematic review relied on published findings only. A substantial proportion of trials with AD do not get published [44] and by doing this we may have missed contributions that demonstrated the superiority of newer AD over BDZ. However, this is an unlikely possibility: it
is the trial that shows the superiority of BDZ over newer AD which is unlikely to be published because there are no major financial incentives (patents) for BDZ [45].

A major drive in the shift from BDZ to AD in anxiety disorders was the risk of dependence with BDZ [46, 47]. However, in due course after their introduction, similar if not more pronounced problems occurred with most of the newer AD [48–51]. Withdrawal reactions and post-withdrawal syndromes may ensue, despite slow tapering, with both types of drugs. While there is a controlled comparison of different types of drugs of the same class (for instance, paroxetine, sertraline and fluoxetine), to the best of our knowledge there is no such comparison between BDZ and AD. Our research group performed a study on 16 patients on BDZ who had recovered from panic disorder upon exposure treatment and had drug treatment slowly tapered and discontinued under optimal conditions; 13 of the 16 patients (81%) reported a withdrawal reaction [52]. A subsequent study used the same methodology in a similar patient population treated with new AD [53]; 9 of the 20 patients (45%) experienced a discontinuation syndrome according to specific criteria. Even though these data cannot substitute a controlled comparison, they indicate that both classes of drugs present with the same type of problem once a discontinuation is attempted with the best possible strategy.

Overall, in GAD, complex phobias and mixed anxiety-depressive disorders, BDZ were better tolerated than both TCA and newer AD, leading to fewer dropouts and adverse reactions. This was also confirmed to occur in panic disorder by the results of our meta-analysis. As to long-term effects, Nardi et al. [39], at a 3-year follow-up of continued monotherapy with either clonazepam or paroxetine in panic disorder, showed that not only was long-term treatment with clonazepam still better in terms of clinical improvement than paroxetine, but it also led to significantly fewer AE. More specifically, during long-term treatment, participants taking paroxetine experienced sexual dysfunctions, drowsiness/fatigue, memory/concentration problems and insomnia more frequently than those treated with clonazepam. The better tolerability found with BDZ compared to AD may be explained by considering a number of clinical data. There is emerging awareness of serious and bothersome side effects that may ensue with long-term treatment with SSRI, such as high rates of sexual dysfunction, bleeding (in particular gastrointestinal), weight gain, risk of fracture and osteoporosis, and hyponatremia [54]. Further, anxiety disorders frequently occur in the setting of medical diseases [55]. An issue that is frequently underestimated is the potential for drug interactions of AD, with special reference to the SSRI medications [56]. Both in terms of long-term treatment side effects and potential for drug interactions, BDZ appear to be much safer than AD [46, 47, 57]. It may well be that the only potential advantage of SSRI versus BDZ is represented by a lower impairment in cognitive and psychomotor skills [58]. Moreover, an additional area of concern regarding long-term treatment with AD drugs in anxiety disorders has emerged: AD may precipitate hypomania and mania also in patients with anxiety disorders [59, 60]. This phenomenon may encompass subsyndromal manifestations [61] and is particularly accentuated in younger patients [62, 63].

Recently, the concept of iatrogenic comorbidity has been introduced [43]. It refers to the lasting effects that previous treatments may entail on the course and responsiveness of illness, such as affective liability [51, 64] and generalized unresponsiveness [65, 66] after AD use. The choice of a specific drug treatment may thus take into consideration issues related to iatrogenic comorbidity.

The findings of this systematic review and the available literature thus lend no support to the shift in the prescribing pattern favoring newer antidepressants over BDZ in the treatment of anxiety disorders. Berney et al. [8] deserve credit for raising this issue in their previous review. Indeed, a reassessment of the use of BDZ is warranted.

**Acknowledgment**

Dr. Offidani’s work was supported in part by a grant from Fondazione Cassa di Risparmio di Cesena awarded to Dr. Tomba.

**Disclosure Statement**

No author or immediate family member has financial arrangements that might represent potential conflicts of interest for the findings presented.

**References**


Benzodiazepines versus Antidepressants in Anxiety Disorders

Psychother Psychosom 2013;82:355–362
DOE: 10.1159/000353198


Begg C: Publication bias; in Cooper H, Br J Psychiatry 1999;13:115–118.


