Chapter 9

Influence of baseline severity on antidepressant efficacy for anxiety disorders: meta-analysis and meta-regression

Ymkje Anna de Vries, Peter de Jonge, Edwin van den Heuvel,
Erick H. Turner, Annelieke M. Roest

British Journal of Psychiatry (2016), 208, 515 - 521

Abstract

Background: Antidepressants are established first-line treatments for anxiety disorders, but it is not clear whether they are equally effective across the severity range.

Aims: To examine the influence of baseline severity of anxiety on antidepressant efficacy for generalized anxiety disorder (GAD), social anxiety disorder (SAD), obsessive-compulsive disorder (OCD), post-traumatic stress disorder (PTSD) and panic disorder (PD).

Methods: Fifty-six trials of second-generation antidepressants for the short-term treatment for an anxiety disorder were included. Baseline and change scores were extracted for placebo and treatment groups in each trial. Mixed-effects meta-regression was used to investigate the effects of treatment group, baseline severity, and their interaction.

Results: Increasing baseline severity did not predict greater improvement in drug groups compared to placebo groups. Standardized regression coefficients of the interaction term between baseline severity and treatment group were 0.04 (95% confidence interval -0.13 to 0.20, p=0.65) for GAD, -0.06 (-0.20 to 0.09, p=0.43) for SAD, 0.04 (-0.07 to 0.16, p=0.46) for OCD, 0.16 (-0.22 to 0.53, p=0.37) for PTSD, and 0.002 (-0.10 to 0.10, p=0.96) for PD. For OCD, baseline severity did predict improvement in both placebo and drug groups equally ($\beta = 0.11$, 95% confidence interval 0.05 to 0.17, p=0.001).

Conclusions: No relationship between baseline severity and the drug-placebo difference was found for anxiety disorders. These results suggest that if the efficacy of antidepressants is considered clinically relevant, they may be prescribed to anxious patients regardless of symptom severity.

Introduction

Anxiety disorders are the most common mental disorders, with a combined 12-month prevalence of 18.1% [4] and lifetime prevalence of 28.8% in the US [1]. Due to their high prevalence, combined with an often early onset and chronic course [1, 3, 373], anxiety disorders are the second most important cause of disability worldwide within the group of mental and behavioral disorders [316].

Antidepressants, including the second-generation selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs), have been found to be efficacious in the treatment of most anxiety disorders, including generalized anxiety disorder (GAD) [116], social anxiety disorder (SAD) [374], obsessive-compulsive disorder (OCD) [123], post-traumatic stress disorder (PTSD) [122], and panic disorder (PD) [119].

However, research in major depressive disorder (MDD) has suggested that the efficacy of antidepressants depends upon the baseline severity of depression. One meta-analysis of the Food and Drug Administration (FDA) database of randomized controlled trials (RCTs) of second-generation antidepressants showed that trials with higher mean baseline severity scores were more likely to be positive [70], while another found a significant interaction between baseline severity and treatment group in predicting improvement, such that the drug-placebo difference is relatively small at low levels of initial severity [69]. Similarly, an additional meta-analysis found no statistically significant difference between antidepressants and placebo in patients with subthreshold (minor) depression [375]. Among 4 analyses using individual patient data, 3 found that baseline severity of depression is associated with antidepressant efficacy [71, 72, 376], while another recent, large study did not find a significant association [76].

Much less is known about the relationship between baseline severity and antidepressant efficacy in the context of anxiety disorders. For OCD, a meta-analysis of 24 antidepressant RCTs found that baseline severity predicted greater improvement in both placebo and drug groups, but there was no evidence for an interaction between baseline severity and treatment group [78]. Recently, a meta-analysis of 12 RCTs of paroxetine for GAD and PD found no evidence for an interaction effect [77]. However, pooling trials for different disorders may obscure differences between disorders; it also necessitated the use of a secondary outcome for PD, rather than the primary, panic-specific outcome in these trials. Other single trials and small pooled analyses that have examined this question have reached contradictory conclusions [377, 378, 379].

The evidence to date, therefore, is (perhaps with the exception of OCD) conflicting, minimal, or absent altogether. To our knowledge, no study has comprehensively investigated whether baseline severity predicts antidepressant efficacy in all anxiety disorders. If antidepressant efficacy does depend upon baseline anxiety severity, this has important consequences for the continued development of guidelines for the treatment of anxiety disorders. Therefore, in order to investigate this question, we conducted a meta-analysis

and meta-regression of RCTs of SSRIs and SNRIs for the short-term treatment of anxiety disorders submitted to the Food and Drug Administration (FDA).

Methods

Study retrieval and data extraction

We obtained drug approval packages (a.k.a. reviews) for all SSRIs and SNRIs approved by the FDA for the short-term treatment of five anxiety disorders (GAD, SAD, OCD, PTSD and PD) [20]. Reviews were downloaded from the FDA website, when available, or requested from the FDA's Division of Freedom of Information [106].

From the FDA reviews, we extracted data on the duration of the trial, drug dose, number of participants, mean score on the primary outcome measure at baseline and endpoint (with standard error (SE) or standard deviation (SD) if available), and the mean change in the primary outcome measure (with SE/SD if available), for drug and placebo groups separately.

For GAD, the primary outcome measure in all included trials was the change from baseline on the HAM-A; for SAD, it was the change on the LSAS; for OCD, the change on the Yale-Brown Obsessive Compulsive Scale (Y-BOCS); and for PTSD, the change on the Clinician-Administered PTSD Scale, part 2 (CAPS-2).

For PD, the primary outcome for most trials was the number of panic attacks over a 1 – 3 week time frame, dichotomized into 0 (remitted) or ≥ 1 (not remitted); 4 trials, however, used the change in number of panic attacks as a primary outcome measure. Therefore, we also extracted data on the proportion of participants that were free of panic attacks at endpoint (remission rate). Furthermore, baseline scores were reported as the number of panic attacks in the 1, 2 or 3 weeks before baseline; we converted all scores to a 2-week time frame.

Data were extracted from the (modified) intention-to-treat analyses only, which used the last observation carried forward (LOCF) method to handle missing data from dropouts. Data were extracted preferably from the statistical review; however, we gave preference to other documents within the drug approval package (e.g. the medical review or administrative correspondence) if they provided more complete data, e.g. the SE/SD, as well as the mean change. If complete data was not available in the FDA review, we attempted to obtain the missing data from secondary sources (trial registries and published journal articles).

Outliers and missing data

Exploration of mean baseline scores, stratified by disorder, revealed no outliers in the dataset for GAD, SAD, OCD and PTSD. For PD, however, 1 placebo group and 1 drug group (from 2 separate trials) had mean baseline scores more than 2 standard deviations greater than the overall mean baseline score (after log-transformation to normalize the distribution); we therefore excluded these groups (but not other groups within the same trial) from our analysis.

We were able to obtain data on the mean change score and its SE/SD (or the remission rate for PD trials) for 46 out of 56 trials. For 40 trials, the FDA review provided complete data; for 2 trials, we obtained data on the SE/SD from the GlaxoSmithKline trial registry; and for 4 trials, we obtained data on the SE/SD (3 trials) or the remission rate (1 trial) from the matching published journal articles, after verifying that mean baseline and change scores matched those in the corresponding FDA review.

For 10 out of 56 trials, not all required data could be obtained from any source. Information on the SE/SD of the change score was missing for 1 PTSD and 6 OCD trials, while information on the remission rate was missing for 3 PD trials. For all 10 trials, however, information on the mean change score was available. For PTSD and OCD, the change score itself was strongly correlated with its SD in groups without missing data; we therefore imputed these missing SDs based upon the change score, group membership and their interaction. For PD, endpoint score was strongly correlated with the remission rate and we therefore imputed missing remission rates based upon the endpoint score, group membership and their interaction. Imputation was performed separately per disorder in SPSS 20, using multiple imputation with Fully Conditional Specification based upon linear regression in order to create 10 imputed datasets.

Statistical analysis

For our main analysis, we calculated effect sizes separately per treatment group, as a single-group pre-post effect size. This approach allowed us to investigate not only the relationship between baseline severity and antidepressant efficacy (drug-placebo difference), but also its underlying cause (i.e. change in the placebo response versus change in the drug response).

For GAD, SAD, OCD and PTSD, the standardized mean difference (SMD) was first calculated based on the (within-group) change score and its standard deviation, according to the formula $SMD = D/SD_D$, where D signifies the change score and SD_D the standard deviation of the change score (25).

By using this formula, we assume that the correlation between baseline and endpoint scores is 0.5, as the true correlation is unknown. We then applied Hedges' correction for

small sample size, where n indicates the number of participants in a group [380]:

$$g = (1 - \frac{3}{4(n-1)-1}) \times SMD$$

The standard error of Hedges' g was computed as follows [380]:

$$SE = (\frac{1}{n})(\frac{n-1}{n-3})(1 + n \times g^2) - (g^2/(1 - \frac{3}{4(n-1)-1})^2$$

For PD, we selected the remission rate itself as our effect measure.

In order to obtain a single effect size for the antidepressant arms of fixed-dose studies, we used a fixed-effects inverse-variance-weighted model to pool these drug groups into one estimate of effect size with its standard error for GAD, SAD, OCD and PTSD. A pooled remission rate was derived for PD by calculating the sample-size-weighted average of the remission rates in the different dose groups. For all disorders, a pooled baseline score was derived by calculating the sample-size-weighted mean of the baseline scores in the different dose groups. Pooling different dose groups may not be appropriate in the presence of a dose-response relationship, but with the exception of venlafaxine, such relationships usually cannot be demonstrated with second-generation antidepressants [381].

All analyses were performed in Stata 13. We performed meta-analyses using the *metan* command, applying a random-effects (DerSimonian-Laird) model to obtain summary statistics by disorder and group. In order to measure heterogeneity, the *heterogi* module within Stata was used to calculate I^2 and its 95% confidence interval [382].

Meta-regression was then performed separately per disorder using the *metareg* command. All meta-regressions were based on a mixed-effects model, used restricted maximum likelihood (REML) estimation of the residual between-study variance, and included group, baseline severity, and their interaction as predictors. The dependent variable was Hedges' g for GAD, SAD, OCD, and PTSD, and the remission rate for PD. Studies were weighted according to the inverse of their variance. For OCD, PTSD, and PD, meta-regression estimates from the 10 multiply imputed datasets were combined using the *mi* suite of commands in Stata.

Sensitivity analyses

As a secondary analysis, we calculated Hedges' g for the drug-placebo difference directly from the exact p-value for the statistical test performed at endpoint (or alternative methods as required [see reference 20]). The trial baseline severity score was calculated as the sample-size-weighted average of all groups (drug as well as placebo) included in the trial. We performed meta-regressions separately per disorder using the *metareg* com-

mand in Stata. The dependent variable was the drug-placebo difference (Hedges' g) for all disorders, and baseline severity was the only predictor in this analysis.

Additionally, to increase statistical power and improve generalizability of results, we expanded our dataset by including extra trials. We included active comparator arms and trials that were not conducted for the purpose of marketing approval, such as trials of other medications (e.g. antipsychotics) in which the antidepressant was used as an active comparator. Trials were obtained from the most recent meta-analyses examining (pharmacological) treatment of anxiety disorders [123, 187, 383]. As this introduced trials with very small sample sizes, additional heterogeneity, and likely reporting bias, this expanded set of trials was examined as a secondary analysis only.

We included parallel-group, placebo-controlled trials with a similar duration (8 – 16 weeks) as our primary set of trials. Trials with a sample that partly or fully overlapped with a trial included in the primary set were excluded, as were trials that did not use a compatible outcome (i.e., HAM-A, LSAS, Y-BOCS, CAPS, or number of panic attacks as a primary or secondary outcome). We repeated our primary analysis for this expanded set of trials.

Results

Description of included studies

A total of 21 reviews were obtained, including nine (formulations of) SSRIs and SNRIs: escitalopram, duloxetine, fluoxetine, fluvoxamine, fluvoxamine controlled release (CR), paroxetine, paroxetine CR, sertraline and venlafaxine extended release (XR). These reviews comprised a total of 59 RCTs: 11 for GAD, 11 for SAD, 13 for OCD, 7 for PTSD and 17 for PD. However, we excluded 2 trials for SAD and 1 trial for PD, as baseline severity scores were incomparable to the other trials. Specifically, the excluded SAD trials did not use the Liebowitz Social Anxiety Scale (LSAS), while the excluded PD trial failed to distinguish between full and limited-symptom panic attacks.

Consequently, 56 trials were included in this study. We only used data from doses recommended by the FDA; therefore, for paroxetine, we excluded the 20-mg dose in 1 OCD trial and the 10- and 20-mg dose in 1 PD trial, as only doses of 40 mg and higher were judged to be effective for these disorders.

All included trials were short-term, randomized, double-blind and placebo-controlled. Some trials also utilized active comparators, but data from these were not included in our primary analyses. Trial duration was 12 weeks for all SAD and PTSD trials, 8 – 10 weeks for GAD trials, 8 – 16 weeks for OCD trials, and 10 – 12 weeks for PD trials.

Trials included adults of 18 years or older (with the exception of 5 OCD trials that included adolescents together with adults) who met DSM-III-R or DSM-IV criteria for

the anxiety disorder under investigation. The majority of trials (71%) used a flexible-dose design, allowing investigators to titrate the dose according to a subject's response, while the remainder (29%) used a fixed-dose design with patients randomized to one of 2 or 3 different dosage groups.

The 56 trials included a total of 55 placebo groups and 78 drug groups, which were pooled into 56 drug groups. The total number of participants across disorders was 14,710, of whom 6,386 were randomized to placebo and 8,324 were randomized to drug. Baseline severity was generally in the moderate to severe range. Participant numbers and baseline severity scores by disorder and treatment group are shown in Table 9.1. A complete listing of included trials, with baseline and change scores (for GAD, SAD, OCD, and PTSD) or remission rates (for PD), is provided in Tables 9.4 and 9.5.

Table 9.1: *Participant numbers and baseline scores by disorder and group.*

	Participants			Mean baseline score (range)	
Disorder	Placebo	Drug	Total	Placebo	Drug
GAD	1628	2219	3847	23.9 (22.1 – 25.9)	24.0 (22.5 – 26.0)
SAD	1255	1379	2634	85.8 (73.3 – 93.9)	86.7 (78.0 – 95.9)
OCD	976	1583	2559	24.5 (22.6 – 26.3)	24.4 (22.6 – 26.6)
PTSD	829	981	1810	74.3 (72.0 – 78.4)	74.4 (72.0 – 77.4)
PD	1698	2162	3860	10.6 (6.2 – 19.2)	11.5 (6.9 – 17.6)

GAD: generalized anxiety disorder; OCD: obsessive-compulsive disorder; SAD: social anxiety disorder; PD: panic disorder; PTSD: post-traumatic stress disorder. Baseline scores for GAD are based on the Hamilton Anxiety Rating Scale (HAM-A), for SAD on the Liebowitz Social Anxiety Scale (LSAS), for OCD on the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS), for PTSD on the Clinician-Administered PTSD Scale (CAPS-2) and for PD on the number of panic attacks in the 2 weeks before baseline.

Meta-analysis

We performed a meta-analysis of effect sizes, stratified by disorder and group (Table 9.2). The placebo group effect size was smallest for OCD (0.49), followed by SAD (0.65). These placebo effects were 59% and 60%, respectively, of the effects found for the corresponding drug groups. By contrast, for GAD and PTSD, the placebo effect sizes were 1.03 and 0.97, respectively; these effects were 76% and 83%, respectively, of the effects found for the corresponding drug groups. For PD, the placebo remission rate was 45%, which was 76% of the rate found in the drug group. Substantial heterogeneity was present for nearly all groups, with I^2 ranging from 46% for the placebo groups in SAD trials to 89% for the drug groups in PTSD trials, although confidence intervals were generally wide.

Table 9.2: *Meta-analysis of effect sizes*

	Effect size (95% CI)		
Disorder	Placebo	Drug	Difference
GAD	1.03 (0.93 - 1.13)	1.35 (1.24 - 1.46)	0.32 (0.16 - 0.47)
SAD	0.65 (0.57 - 0.74)	1.08 (0.99 - 1.18)	0.43 (0.29 - 0.57)
OCD	0.49 (0.39 - 0.59)	0.83 (0.73 - 0.92)	0.34 (0.19 - 0.48)
PTSD	0.97 (0.81 - 1.13)	1.17 (0.91 - 1.44)	0.20 (-0.15 - 0.55)
PD	0.45 (0.38 - 0.52)	0.59 (0.55 - 0.65)	0.14 (0.06 - 0.22)

GAD: generalized anxiety disorder; OCD: obsessive-compulsive disorder; SAD: social anxiety disorder; PD: panic disorder; PTSD: post-traumatic stress disorder. The effect size is expressed as Hedges' g for GAD, SAD, OCD and PTSD, and as remission rate for PD.

Meta-regression

For both GAD and SAD, neither baseline severity nor the interaction between group and baseline severity were statistically significant predictors of the effect size, although group membership was (p=0.001 and <0.001, respectively) (Table 9.3 and Figure 9.1). For OCD, the interaction between baseline severity and group membership was not significant, but the main effects of both group membership and baseline severity were (group: p<0.001; baseline: p=0.001), indicating that the (positive) slope of the association between baseline severity and effect size was similar in the placebo and drug groups. For PTSD, none of the predictors achieved statistical significance.

For PD, we modeled the relationship between baseline severity and the remission rate. In this model, the interaction between group membership and baseline severity was not significant (Table 9.3 and Figure 9.2). Group membership was a significant predictor of the remission rate (p=0.007), while baseline severity was not.

Paralleling these results, we found that including treatment group in the model reduced between-group heterogeneity for GAD (I^2 decreasing from 83% to 67%), SAD (84% to 49%), OCD (75% to 52%), and PD (86% to 77%), while it did not reduce heterogeneity for PTSD (85% to 84%). Including the main effect of baseline reduced heterogeneity for OCD only (52% to 14%), while including the interaction did not reduce heterogeneity further for any disorder.

Sensitivity analyses

In our secondary analysis, baseline severity was not a significant predictor of the drug-placebo difference for any disorder, and the effect of baseline severity was positive (i.e. in the expected direction) only for PTSD ($\beta = 0.084$, p=0.42). For the other disorders, the effect ranged from -0.064 to -0.002 (p-values ranging from 0.170 to 0.860).

Table 9.3: *Meta-regression analysis*

		Model 1 (with interaction)		Model 2 (without interaction)	
Disorder	Predictor	β (95% CI)	p	β (95% CI)	p
GAD	Group	0.31 (0.15 - 0.47)	0.001	0.32 (0.16 - 0.48)	0.001
	Baseline	-0.03 (-0.15 - 0.09)	0.60	-0.01 (-0.09 - 0.07)	0.77
	G x B	0.04 (-0.13 - 0.20)	0.65		
SAD	Group	0.43 (0.29 - 0.57)	0.001	0.43 (0.29 - 0.56)	0.001
	Baseline	0.06 (-0.04 - 0.17)	0.21	0.04 (-0.03 - 0.11)	0.29
	G x B	-0.06 (-0.20 - 0.09)	0.43		
OCD	Group	0.35 (0.0.23 - 0.46)	0.001	0.35 (0.24 - 0.47)	0.001
	Baseline	0.09 (-0.004 - 0.17)	0.04	0.11 (0.05 - 0.17)	0.001
	G x B	0.04 (-0.07 - 0.16)	0.46		
PTSD	Group	0.20 (-0.17 - 0.56)	0.25	0.20 (-0.16 - 0.56)	0.24
	Baseline	-0.01 (-0.28 - 0.27)	0.96	0.07 (-0.12 - 0.26)	0.41
	G x B	0.16 (-0.22 - 0.53)	0.37		
PD	Group	0.14 (0.05 - 0.24)	0.006	0.14 (0.05 - 0.24)	0.005
	Baseline	-0.01 (-0.11 - 0.09)	0.85	-0.01 (-0.08 - 0.06)	0.81
	G x B	0.002 (-0.10 - 0.10)	0.96		

GAD: generalized anxiety disorder; OCD: obsessive-compulsive disorder; SAD: social anxiety disorder; PD: panic disorder; PTSD: post-traumatic stress disorder. G x B = Group x Baseline interaction.

The expanded set of trials included 9 additional trials for GAD [384, 385, 386, 387, 388, 389, 390, 391, 392], 10 additional trials for SAD [393, 394, 395, 396, 397, 398, 399, 400, 401, 402], 6 additional trials for OCD [403, 404, 405, 406, 407, 408], 6 additional trials for PTSD [409, 410, 411, 412, 413, 414], and 6 additional trials for PD [415, 416, 417, 418, 419, 420]. The range of baseline severity was increased slightly to substantially for all disorders in this expanded set of trials. However, no statistically significant interaction effects were found for any disorder, although the main effect of baseline became statistically significant for GAD, PTSD, and SAD, in addition to OCD (Table 9.6 in the Appendix).

Discussion

Principal findings

We found no evidence that baseline severity of disorder affects the efficacy of second-generation antidepressants in the short-term treatment of anxiety disorders. This finding stands in remarkable contrast to what has been reported in studies investigating MDD.

For OCD, baseline severity did predict change in score in both placebo and drug groups, but no differential effect was apparent. This suggests that patients with more severe OCD may show substantially greater improvement with antidepressant treatment than

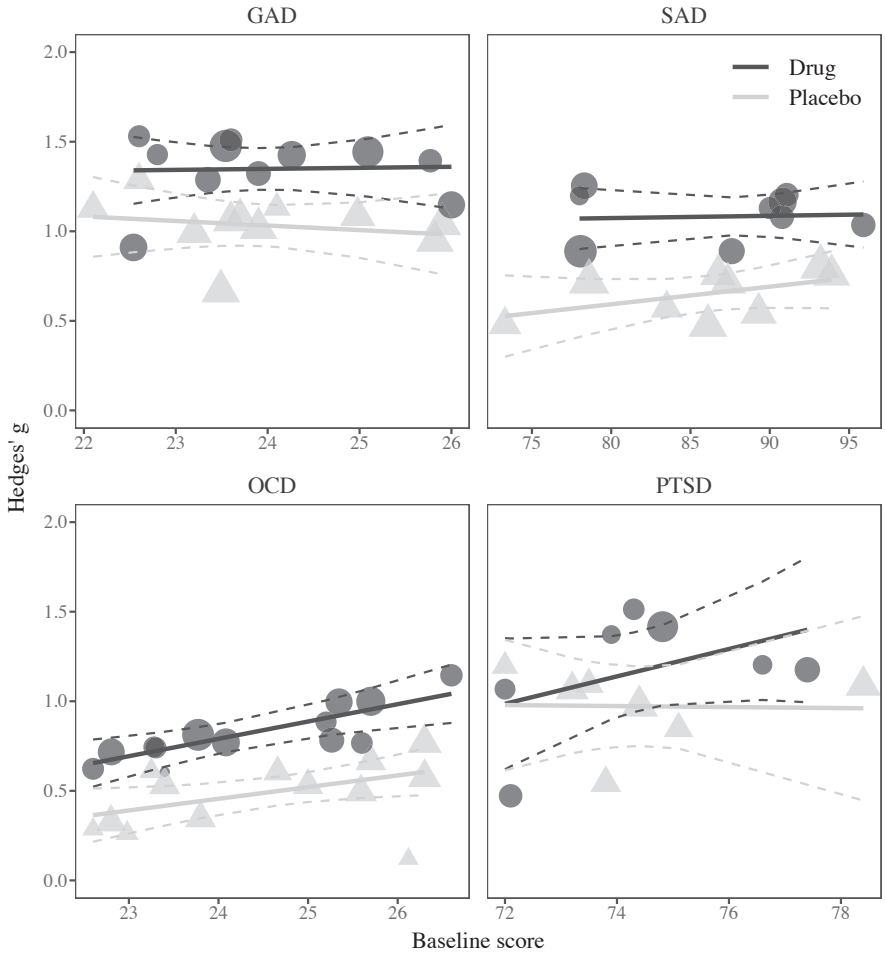

Figure 9.1: *Meta-regression analysis for GAD, SAD, OCD, and PTSD. Data points are sized proportionally to the inverse of their standard error.*

patients with milder OCD, but this is not due to improved antidepressant efficacy in severe OCD. Our sensitivity analysis, in which 37 additional trials were included, suggests that a similar regression to the mean effect might also occur in other anxiety disorders, but confirmed the lack of evidence for an interaction effect in all anxiety disorders.

Effect sizes for the drug-placebo difference were unaffected by inclusion of baseline severity as a main effect or in interaction with group, but they were generally smaller than a criterion for clinical significance previously used (though without clear justification) for MDD (d=0.5) [69]. However, it has been shown that effect sizes exceeding 0.5 are not achieved by most current treatments, either in psychiatry or in general medicine [421].

Furthermore, clinical significance is context- and disorder-specific [422]. An empirically-

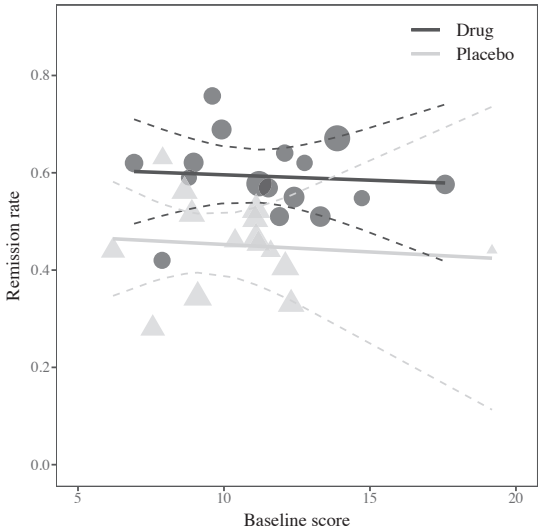

Figure 9.2: *Meta-regression analysis for PD. Data points are sized proportionally to the inverse of their standard error.*

derived criterion of d=0.24 has been suggested to be a more meaningful threshold for clinically significant efficacy for MDD [423]. Although not all included anxiety disorders met that threshold in our primary analysis, the drug-placebo difference may have been underestimated slightly due to our analytical approach. We have previously analyzed the trials included in this study by conventional meta-analytic methods and found effect sizes of 0.27 and greater [20]. Further research is required to establish whether antidepressant efficacy in the different anxiety disorders may be considered clinically significant in all or in subsets of patients.

Additionally, the expected efficacy of medication is only one of many factors that play a role in the decision to prescribe antidepressants for an individual patient. Other issues, such as the expected burden of side effects, the burden posed by the disorder itself, the anticipated course of the disorder, and the availability, acceptability, and efficacy of alternative treatment options like psychotherapy, must also be considered [424]. These considerations might lead to different prescribing decisions for patients with mild versus severe anxiety disorders, even in the absence of differential efficacy.

Comparison with depression

For MDD, it has been suggested that the threshold for "antidepressant-treatable depression" should be higher [425], given that mildly depressed patients do not show a strong differential response to antidepressants over placebo. It is possible that this threshold needs to be even higher in anxiety disorders, and that we would find greater efficacy of

antidepressants in the most severely anxious patients than we do in patients with less extreme anxiety.

Another possible explanation lies in the difference in chronicity between anxiety disorders and MDD. Depressive episodes tend to be relatively short, with a median duration of 3 – 6 months [3, 426], while the median duration of an anxiety episode has been estimated at 16 months [3]. However, duration of a depressive episode is positively correlated with severity [3]. This suggests that, all other things being equal, patients with mild depression are more likely to remit spontaneously within the short time frame of a clinical trial than patients with severe depression.

Since a drug effect cannot be demonstrated in patients who remit spontaneously, efficacy would be expected to be reduced in a patient group with a high likelihood of spontaneous remission. Consequently, the correlation between severity and episode duration may explain the increase in antidepressant efficacy with increasing depression severity. Although severity also correlates with episode duration in anxiety disorders [427], this may play less of a role within the context of a short clinical trial. Given the long median duration of an episode of anxiety, spontaneous remission rates would be expected to be relatively low across the severity range.

Alternatively, we may have been unable to detect a relationship between baseline severity and antidepressant efficacy due to a restriction in range. As trials generally have a minimum severity requirement, as well as exclusion criteria that tend to exclude severely ill patients (e.g., regarding treatments received and comorbidity), baseline severity was restricted to the moderate to very severe range in our primary analyses, although the range was extended somewhat in our sensitivity analyses. However, previous studies in MDD, which were able to detect a relationship between baseline severity and antidepressant efficacy, had similarly restricted severity ranges.

We may also have had insufficient power to find a statistically significant interaction: although the total number of trials included in the current study was large, the number of trials per disorder was limited to 7 – 16. However, the results do not suggest a meaningful trend toward increasing efficacy with increasing severity, except possibly for PTSD, for which we had the fewest available trials. The estimates of the interaction effect were substantially lower (≤ 0.04, except for PTSD) or even in the opposite direction as the estimate previously reported for depression [69]. These estimates are small enough to be of limited clinical relevance, although some confidence intervals encompass values that may be considered clinically relevant.

Furthermore, our secondary analyses, which included 6 to 10 additional trials per disorder, also showed no evidence for increasing efficacy with increasing severity, even for PTSD. Larger samples would be required to conclusively exclude the possibility of even small interaction effects, but unfortunately the number of randomized trials that have been conducted is limited, and consequently a larger sample of trials is not available.

Strengths and limitations

Among the strengths of the current study is the fact that we were able to obtain a dataset that is free from the influence of publication or outcome reporting bias. An additional strength of the study is the high quality of the included trials. As these trials were conducted for the purpose of obtaining marketing approval, they were required to meet strict standards on internal validity (blinding, randomization, etc.). This also ensured that trials conducted for the treatment of the same disorder were, in general, quite comparable.

A limitation is that we were forced to use a different outcome measure for PD (remission) than we did for the other disorders; consequently, we cannot rule out the possibility that an interaction might have been found if we had used a continuous outcome (such as change) instead.

An additional, important limitation is that we did not have access to individual patient data and hence used summary data instead. Use of summary data necessarily causes a loss of information, due to averaging out inter-individual variation within trials, and a concomitant loss of power. Although individual patient data have historically been very difficult to obtain, novel initiatives are now beginning to increase their accessibility to researchers (e.g. clinicalstudydatarequest.com).

Given our current results, which suggest that the influence of baseline severity on antidepressant efficacy may be different for anxiety disorders than for depression, future research making use of individual patient data is essential to provide a definitive answer to this important question.

Conclusions

In conclusion, we found no evidence for an interaction between treatment group and baseline severity in predicting symptom change. It has previously been recommended that treatment with antidepressants should be withheld for mild depression [425, 428]. Our results show that this cannot be simply extrapolated to anxiety disorders, and it would therefore be premature to recommend that antidepressants be withheld for mild anxiety. What defines a clinically relevant effect size remains a matter of debate, but if the effect of antidepressants on anxiety is considered clinically relevant, these results suggest that antidepressants may be prescribed to anxious patients regardless of symptom severity.

Appendix

Table 9.4: *Characteristics of included trials for GAD, SAD, OCD and PTSD*

			Placebo group			Antidepressant group		
Disorder	Drug	Trial	N	Baseline	Change (SD)	N	Baseline	Change (SD)
GAD	Escitalopram	MD-05 [130]	128	22.1	7.7 (6.8)	124	22.8	9.6 (6.7)
		MD-06 [130]	138	22.6	7.6 (5.9)	143	22.6	9.2 (6.0)
		MD-07 [131]	153	23.2	7.4 (7.4)	154	23.6	11.3 (7.4)
	Paroxetine	641 [132]	180	23.9	9.6 (9.4)	188	23.8	12.5 (8.2)
						197	23.3	12.2 (8.4)
		642 [133]	163	23.6	9.5 (8.9)	161	23.9	11.8 (8.9)
		637 (N/A)	183	25.9	11.3 (10.8)	181	26.0	12.4 (10.8)
	Duloxetine	HMBR [134]	173	25.8	8.4 (8.8)	165	25.1	12.8 (8.7)
						169	25.1	12.5 (8.7)
		HMDT [135]	158	23.5	5.9 (8.8)	161	22.5	8.1 (8.9)
		HMDU [136]	158	25.0	9.2 (8.4)	149	25.8	11.8 (8.4)
	Venlafaxine XR	210 [137]	96	24.1	9.5 (8.3)	86	24.7	11.1 (8.8)
						81	24.5	11.7 (7.8)
						86	23.6	12.1 (7.5)
		214 [138]	98	23.7	8 (7.2)	87	23.7	10.6 (7.6)
						87	23.0	9.8 (8.0)
SAD	Fluvoxamine	3107 [149]	125	89.3	13.2 (24.1)	110	90.0	26.6 (23.4)
		3108 [150]	148	93.9	26.2 (34.4)	126	95.9	34.6 (33.2)
	Paroxetine	502 [151]	145	86.1	15.6 (32.8)	136	87.6	29.4 (32.9)
		382 [152]	92	83.5	14.5 (25.2)	90	78	30.5 (25.2)
		454 [153]	92	73.3	15.0 (31.1)	89	79.8	31.4 (29.5)
						88	77.5	24.5 (30.3)
						91	76.9	25.2 (30.0)
	Paroxetine CR	790 [154]	184	78.6	17.6 (24.4)	185	78.3	31 (24.6)
	Sertraline	R-0601 [155]	196	93.2	21.4 (26.6)	205	90.8	31.3 (26.8)
	Venlafaxine XR	387 [158]	138	86.8	19.9 (26.1)	133	91.1	31 (25.6)
		393 [159]	135	87.4	22.1 (30.9)	126	90.8	32.8 (30.2)
OCD	Fluoxetine	HCEP 1 [165]	47	23.0	1.2 (4.5)	47	22.9	5.5 (7.1)
						45	22.4	4.3 (5.3)
						47	23.1	4.2 (6.7)
		HCEP 2 [165]	41	26.1	0.6 (4.6)	39	24.4	3.5 (5.9)
						41	25.4	6.9 (8.1)
						42	26.0	9.1 (9.4)
		E079 [166]	56	23.3	3.7 (6.0)	52	23.8	5.1 (6.4)
						52	25.5	4.7 (6.9)
						54	23.0	6.1 (6.9)
	Fluvoxamine	5529 (N/A)	80	22.8	1.7 (?)	79	23.3	4.9 (?)
		5534 [167]	77	23.8	1.7 (4.9)	78	22.6	4.0 (6.3)

continued

Table 9.4: *Characteristics of included trials for GAD, SAD, OCD and PTSD*

			Placebo group			Antidepressant group		
Disorder	Drug	Trial	N	Baseline	Change (SD)	N	Baseline	Change (SD)
	Fluvoxamine CR	3103 [168]	119	26.3	5.9 (7.6)	113	26.6	8.7 (7.5)
	Paroxetine	116 [169]	88	25.6	3.4 (6.8)	83	25.4	6.3 (6.7)
						83	25.3	7.3 (6.7)
		118 (N/A)	75	24.7	4.6 (7.5)	79	23.3	5.6 (7.5)
		136 [170]	99	26.3	3.9 (?)	198	25.7	6.9 (?)
	Sertraline	237/248 [171]	44	22.6	1.5 (?)	43	23.4	3.8 (?)
		371/372 [172]	84	23.4	3.4 (?)	79	23.2	6.0 (?)
						81	24.6	4.5 (?)
						80	23.5	6.2 (?)
		546 [173]	79	25	3.6 (?)	85	25.2	6.5 (?)
		495 (N/A)	87	25.7	5.0 (?)	83	25.6	5.4 (?)
PTSD	Sertraline	641 [160]	82	73.8	15.4 (28.1)	84	72.1	13.1 (27.5)
		682 (N/A)	94	72.0	27.9 (?)	94	72.0	27.4 (?)
		640 [161]	104	73.5	26.2 (23.8)	98	73.9	33 (23.9)
		671 [162]	90	75.1	23.2 (27.1)	93	76.6	33 (27.2)
	Paroxetine	651 [163]	167	74.4	25.3 (25.8)	166	75.3	39.6 (25.8)
						156	74.3	37.9 (28.7)
		648 [164]	133	73.2	24.7 (23.1)	136	74.3	35.5 (23.3)
		627 (N/A)	159	78.4	26.2 (24.0)	154	77.4	30.8 (26.1)

CR: controlled release; GAD: generalized anxiety disorder; OCD: obsessive-compulsive disorder; SAD: social anxiety disorder; PTSD: post-traumatic stress disorder; XR: extended release.

Table 9.5: *Characteristics of included trials for panic disorder*

		Placebo group			Antidepressant group		
Drug	Trial	N	Baseline	Remission rate	N	Baseline	Remission rate
Fluoxetine	HCJC [429]	90	7.6	28	90	7.9	42
	HCJB (N/A)	104	6.2	44	107	6.9	62
Paroxetine	120 [139]	69	11.6	43.9	72	9.6	75.8
	187 [141]	123	12.3	33	123	11.9	51
	223 (N/A)	68	7.9	63	77	8.8	59
Paroxetine CR	494 [142]	129	11.1	50.4	122	9.9	68.9
	495 [142]	136	8.9	51.5	123	11.5	56.9
	497 [142]	130	8.7	56.2	132	9.0	62.1
Sertraline	629 [143]	87	10.4	46	79	12.8	62
	630 [144]	88	11.2	?	88	12.1	?
	529 [129]	-	-	-	42	20.3	?
					41	21.3	?
					44	11.5	?
	514 (N/A)	38	19.2	?	38	14.1	?
					36	15.4	?
Venlafaxine XR	398 [145]	156	9.1	34.4	158	11.0	54.1
					159	11.4	61.4
	399 [146]	157	11.1	46.5	156	15.7	64.1
					160	12.1	70.0
	353 [147]	155	12.1	40.6	155	13.3	51.0
	391 [148]	168	11.1	52.4	160	12.4	55.0

CR: controlled release; XR: extended release.

Table 9.6: *Secondary meta-regression analysis with expanded set of trials*

		Model 1 (with interaction)		Model 2 (without interaction)	
Disorder	Predictor	β (95% CI)	p	β (95% CI)	p
GAD	Group	0.35 (0.17, 0.52)	<0.001	0.34 (0.17, 0.51)	<0.001
	Baseline	0.11 (-0.03, 0.25)	0.13	0.13 (0.04, 0.23)	0.007
	G x B	0.05 (-0.13, 0.23)	0.57		
SAD	Group	0.47 (0.35, 0.59)	<0.001	0.45 (0.33, 0.58)	<0.001
	Baseline	0.15 (0.04, 0.25)	0.011	0.10 (0.02, 0.18)	0.012
	G x B	-0.08 (-0.23, 0.07)	0.29		
OCD	Group	0.35 (0.23, 0.47)	<0.001	0.36 (0.23, 0.48)	<0.001
	Baseline	0.14 (0.03, 0.25)	0.013	0.18 (0.10, 0.27)	<0.001
	G x B	0.09 (-0.08, 0.25)	0.28		
PTSD	Group	0.21 (-0.03, 0.45)	0.08	0.21 (-0.03, 0.44)	0.08
	Baseline	0.17 (-0.06, 0.40)	0.14	0.20 (0.05, 0.35)	0.013
	G x B	0.06 (-0.26, 0.37)	0.71		
PD	Group	0.13 (0.06, 0.20)	0.001	0.13 (0.06, 0.20)	0.001
	Baseline	0.00 (-0.07, 0.07)	0.97	-0.03 (-0.07, 0.02)	0.20
	G x B	-0.05 (-0.13, 0.04)	0.30		

GAD: generalized anxiety disorder; OCD: obsessive-compulsive disorder; SAD: social anxiety disorder; PD: panic disorder; PTSD: post-traumatic stress disorder. G x B = Group x Baseline interaction.