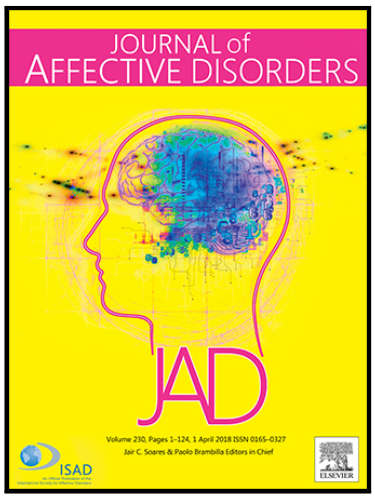

The Magnitude and Heterogeneity of Antidepressant Response in Depression: A Meta-analysis of over 45,000 patients

Xin Guo PhD , Robert A. McCutcheon MRCPsych , Toby Pillinger MRCP , Yuya Mizuno PhD , Sridhar Natesan PhD , Kirsten Brown , Oliver Howes PhD

PII: S0165-0327(20)32547-7
DOI: https://doi.org/10.1016/j.jad.2020.07.102
Reference: JAD 12272

To appear in: *Journal of Affective Disorders*

Received date: 24 March 2020
Revised date: 1 July 2020
Accepted date: 6 July 2020

Please cite this article as: Xin Guo PhD , Robert A. McCutcheon MRCPsych , Toby Pillinger MRCP , Yuya Mizuno PhD , Sridhar Natesan PhD , Kirsten Brown , Oliver Howes PhD , The Magnitude and Heterogeneity of Antidepressant Response in Depression: A Meta-analysis of over 45,000 patients, *Journal of Affective Disorders* (2020), doi: https://doi.org/10.1016/j.jad.2020.07.102

Highlights

- **Question**: Is there greater heterogeneity in response to antidepressant treatment for major depression relative to placebo?
- **Findings**: Compared with placebo, antidepressant-treated patients with MDD demonstrated greater magnitude and lower variability of change of symptom severity.
- **Meaning**: Our findings indicate that improvement with antidepressant treatment is less variable and greater than placebo treatment. This is not consistent with our hypothesis that there are distinct sub-groups of treatment responsive and resistant patients with major depression, our results instead suggest that antidepressants show a relatively uniform effect.

The Magnitude and Heterogeneity of Antidepressant Response in Depression: A Meta-analysis of over 45,000 patients

Abbreviated title**: Heterogeneity Meta-analysis**

Xin Guo, PhD[1,2*], Robert A. McCutcheon, MRCPsych[1,3*], Toby Pillinger, MRCP[1,3], Yuya Mizuno, PhD[1,3,5], Sridhar Natesan, PhD[1], Kirsten Brown[1,3], Oliver Howes, PhD[1,3,4]

[1] Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

[2] Department of Psychiatry, Renmin Hospital of Wuhan university, Wuhan, China

[3] Psychiatric Imaging Group, MRC London Institute of Medical Sciences, Hammersmith Hospital, London, UK

[4] Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London, UK

[5] Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan

* These authors contributed equally

Corresponding Author

Oliver D. Howes, MRCPsych, PhD

Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience,

King's College London

16 De Crespigny Park, London, SE5 8AF, UK

Phone +44 (0)207 848 0080

Email: oliver.howes@kcl.ac.uk

Abstract

OBJECTIVE: To determine the relative variability and magnitude of symptomatic improvement in antidepressant-treated individuals compared to placebo-treated individuals, and to investigate moderating factors.

METHODS: Multiple database and previous publications were searched through February 2019 to identify all randomized controlled trials comparing placebo and antidepressants in acute treatment of depression. Primary outcome was relative variability of change in symptom severity in antidepressant-treated individuals compared to placebo-treated patients quantified using the coefficient of variation ratio (CVR).

RESULTS: Of 9,389 identified records, 134 were found to be eligible (total n=46,646). Antidepressant-treated patients showed a significantly greater magnitude (g=0.28, 95% CI 0.25-0.30, p<.0001) and lower variability (CVR = 0.94, 95% CI 0.93-0.95, p<.0001) of change in symptom severity relative to placebo-treated patients. Antidepressant-related improvement was more uniform in older studies (z=3.01, p=.003) and in studies where antidepressants showed greater efficacy (z=-7.21, p<.0001). Imipramine, moclobemide, amitriptyline and mirtazapine showed significantly lower CVR than several other

antidepressants. However, there's no difference in CVR between multiple/single-neurotransmitter profile antidepressants (z=-0.01, p=.99). Variability in placebo response has reduced in more recent studies (z=-4.78, p<.0001) and studies including severer patients (z=-2.26, p=.02).

CONCLUSION: There is lower variability and greater magnitude of change in symptom severity with antidepressant treatment relative to placebo. This is not consistent with our hypothesis that there are distinct sub-groups of treatment-responsive and treatment-resistant patients with major depression. Our results instead suggest that antidepressants show a relatively uniform effect.

1. Introduction

Major depressive disorder (MDD) is a leading cause of disability and healthcare costs worldwide. (Chisholm et al., 2016; Friedrich, 2017; WHO, 2017) Meta-analysis of randomized clinical trial (RCTs) has demonstrated that antidepressants are more effective than placebo in adults with MDD. (Fournier et al., 2010a) However, response to antidepressants is thought to be heterogeneous, and 50-60% of patients display an inadequate response to initial antidepressant treatment. (Otte et al., 2016; Turner et al., 2008) It has been proposed that major depression consists of treatment-resistant and responsive sub-types, and that around a third of individuals diagnosed with depression have a treatment resistant form of the illness,(Rush et al., 2006) which is generally defined as an inadequate response to at least two adequate trials of different classes of antidepressants. (Conway et al., 2017)

If, underlying these clinical criteria, there are distinct antidepressant treatment-responsive and resistant sub-groups of patients, this suggests that in clinical trials there

will be a greater variability of change in symptom severity seen in the active arm compared to the placebo arm. This is because the active arm will contain patients who benefit from treatment (in terms of the change in symptom severity) to a greater extent than if they had been allocated a placebo, and those that show no improvement beyond that which would have occurred if receiving placebo, whereas the placebo group will show only change associated with the latter. Figure 1 illustrates how individual differences in change of symptom severity to active drug treatment compared to placebo may affect group level distributions. In Figure 1A there is a uniform effect of active treatment over and above the effect of placebo treatment and this results in a greater mean difference but no increase in variability in the drug treated arm. However, in the scenarios illustrated in Figure 1B and 1C there is a non-uniform effect of active treatment and this results in increased variability in the distribution at group level in the active treatment arm compared to the placebo arm. Figure 1C illustrates the expected effect on the distribution of symptom change under the hypothesis that there is a substantial sub-group of patients who are antidepressant treatment resistant (showing a response similar to the placebo group) and a sub-group who show a specific antidepressant treatment effect. Thus, a key prediction of the treatment responsive and resistant subtype hypothesis is that there will be greater variability of change in

symptom severity in patients assigned to active treatment compared to those assigned to placebo treatment.

Advances in meta-analytic techniques mean that in addition to calculating differences in the standardized mean differences between treatments, it is now possible to integrate information from multiple studies in order to quantify the relative magnitude of difference in variability between two treatments to test these predictions. This uses the coefficient of variability ratio (CVR), which is a measure of variability of symptom change between patients treated with antidepressants and placebo while accounting for the difference in mean change. (Brugger et al., 2017; McCutcheon et al., 2019; Nakagawa et al., 2015; Winkelbeiner et al., 2019) In addition, given evidence that placebo response is increasing and that trial characteristics have been found to influence the magnitude of change of symptom severity, we sought to investigate if these factors also influence the variability in change of symptom severity. Finally, we sought to test if variability in change of symptom severity is different between patients receiving an antidepressant that acts on multiple neurotransmitter systems (a multiple neurotransmitter profile drug) relative to those receiving one that acts on a single neurotransmitter system (a single neurotransmitter profile drug). This rests on the assumptions that there are

differences amongst patients in the mechanism required for a treatment effect,(Willner et al., 2013) and that a multiple neurotransmitter profile drug targets more of these than a single neurotransmitter profile drug, potentially obscuring subgroup phenomena and resulting in a more uniform change of symptom severity across the treated sample. However, it should be appreciated that these assumptions may not hold, for example if all drugs act through a final common pathway, in which case there would be no difference in variability between multiple and single neurotransmitter profile drugs.

2. Methods

2.1. Search Strategy and Study Selection

We followed PRISMA guidelines (PRISMA Checklist in Supplement), and registered the study on PROSPERO (CRD42019141155). (Liberati et al., 2009) We searched the MEDLINE, PsycINFO and EMBASE, and Cochrane Central Register of Controlled Trials, Clinicaltrials.gov, clinicaltrialsregister.eu, accessdata.fda.gov, from the data of their inception to February 1, 2019, for randomized controlled trials (RCTs) that have been published in the English language, and checked references of included studies and previous reviews. Our search strategy was based on previously conducted reviews,

(Cipriani et al., 2018) and full search terms are in Supplement. In brief we searched for RCTs of one or more of the following antidepressants: (agomelatine, amineptine, amitriptyline, amoxapine, bupropion, citalopram, clomipramine, desipramine, desvenlafaxine, dosulepin, doxepin, duloxetine, fluoxetine, fluvoxamine, imipramine, isocarboxazid, lofepramine, mianserin, maprotiline, mirtazapine, moclobemide, nefazodone, nortriptyline, paroxetine, phenelzine, protriptyline, reboxetine, selegiline, sertraline, tranylcypromine, trazodone, trimipramine, venlafaxine, vilazodone, vortioxetine, escitalopram) AND placebo AND (Depress$ OR dysthymi$ OR mood$ OR affecti$ OR anhedoni$ OR dysphori$) AND (random* or clinic* trial). We included RCTs comparing antidepressants with placebo or another active antidepressant as oral monotherapy for the acute treatment of adults with a primary diagnosis of major depressive disorder according to standard, operationalized diagnostic criteria (Feighner criteria, Research Diagnositic Creteria, DSM-III, DSM-III-R, DSM-IV, DSM-5, and ICD-10). (Cipriani et al., 2018) We also contacted study authors and drug manufacturers to supplement incomplete reports of the original papers or provide data for unpublished studies. Screening and selection of studies was performed independently by three of the authors (X.G., R.M., & K.B.), with each study assessed by a minimum of two researchers. Disagreements were resolved via discussion with R.M.

Inclusion criteria consisted of 1) randomized, double-blind, placebo-controlled trials, 2) active arms consisting of monotherapy with an antidepressant licensed for the treatment of depression, 3) adults aged 18-65 with an acute exacerbation of major depression. Exclusion criteria included 1) grey literature (with the exception of clinical study reports), 2) studies specifically investigating antidepressants in individuals with severe physical or psychiatric comorbidities (see more details in eMethods in Supplement), 3) relapse prevention studies, 4) trials of intravenous or injectable forms of medication and, 5) trials of medication combination or augmentation strategies. Benzodiazepine and/or hypnotics used as part of standard clinical treatment were allowed. Neuromodulation was not allowed. For psychotherapy, simple suggestions such as advising increased physical activity were included. Specific organized systems of psychotherapy (for example, cognitive behavioral therapy) were not allowed.

2.2. Data Extraction and Processing

We extracted both the mean and variance (standard deviation (SD), standard error, or confidence intervals) of symptom change for depression symptoms ratings from each

study for the active and placebo groups, as measured using the Hamilton Rating Scale for Depression 17/21/24 items (HDRS-17/ HDRS-21/ HDRS-24), (Williams, 1988) Montgomery-Asberg Depression Scale (MADRS).(Montgomery et al., 1979) For total symptoms in studies reporting outcomes using both the HDRS-17/21/24 and MADRS, the HDRS-17 was used in our analysis to ensure consistency in outcome measures and because it was the most commonly reported measure. To make the baseline severity comparable across studies, we converted all values into HDRS-17 (eMethods in Supplement). If a study contained multiple active treatment arms, the number of patients in the placebo group was divided by the number of arms. (Higgins et al., 2008) If information was missing, we contacted authors to request the missing information. We also extracted year of publication, baseline symptom severity, study duration, the antidepressant used, antidepressant dose in fluoxetine equivalent, (Bollini et al., 1999; Hayasaka et al., 2015; Jakubovski et al., 2015) participant age and gender. The quality of included studies was rated using the Cochrane Risk of Bias Assessment Tool (see further detailed criteria for risk of bias in eMethods in Supplement). (Higgins et al., 2011) Discrepancies were resolved by consensus and discussion with the other authors.

2.3. Variability Outcome Measures

The relative variability between changes of symptom severity to an antidepressant and a placebo treatment in depression patients can be quantified using the log variability ratio (VR):

$$VR = ln\left(\frac{\delta_a}{\delta_p}\right) = ln\left(\frac{S_a}{S_p}\right) + \frac{1}{2(n_a - 1)} - \frac{1}{2(n_p - 1)}$$

Where δ_a and δ_p are the unbiased estimates of the population standard deviation for the change in symptom severity rating score of the antidepressant treated, and placebo treated groups respectively. s_a and s_p are the reported SD_s, while n_a and n_p are the sample sizes.

As noted for a range of biological systems, and as we have previously found for other psychiatric conditions and treatments, variance shows a strong association with the mean. (Brugger et al., 2017; McCutcheon et al., 2019; Winkelbeiner et al., 2019) This can be accounted for using the log coefficient of variation ratio (CVR), which adjusts the lnVR for mean differences between groups:

$$CVR = ln\left(\frac{\delta_a/\bar{x}_a}{\delta_p/\bar{x}_p}\right) = ln\left(\frac{S_a/\bar{x}_a}{S_p/\bar{x}_p}\right) + \frac{1}{2(n_a - 1)} - \frac{1}{2(n_p - 1)}$$

Where $\overline{x}_a$ and $\overline{x}_p$ are the mean symptom scores for antidepressant and placebo treated groups respectively. The use of CVR to quantify group differences in variability is appropriate only where the data has a true zero point, (Senn, 2018) and this is not appropriate for raw change scores as they can be positive or negative. We therefore converted values of mean change into adjusted mean change of symptom severity (eMethods in Supplement). A value of CVR below 1 indicates greater variability in the placebo arm, while a value above 1 indicates greater variability in the antidepressant arm. CVR was used as the primary outcome, as otherwise differences in variability can primarily reflect differences in mean, although the VR results are also presented as a secondary analysis.

2.4. Statistical Analysis

All statistical analyses were carried out in the statistical programming language R (version-3.6.0), (Team, 2013) primarily using the "metafor" package (version-2.0.0). (Viechtbauer, 2010) The correlation coefficient between studies' mean and SD, was weighted by study size with the "weights" package (version-1.0), while plots were generated using "ggplot2" (version-2.2.1).

Our primary outcome measure was CVR and, to test our main hypothesis that there was greater variability in the antidepressant treatment group, we grouped all antidepressants together and compared them to placebo. We also analyzed each antidepressant separately to determine if there were drug specific effects, these comparisons were based on indirect comparisons (comparisons of treatment variability between antidepressants is based on the comparisons of variability of antidepressants and placebo treatment). To test the hypothesis that multiple neurotransmitter profile antidepressants would be associated with less variability in change of symptom severity, we compared CVR between classes of antidepressants based on Neuroscience based Nomenclature (NbN)(Uchida et al., 2016) (eMethods in Supplement). In addition to the variability measures we also calculated the standardized mean difference (Hedges' g) between placebo and antidepressant treated groups using a random effects model.

Subgroup analyses were performed by performing separate meta-analyses for studies that used a placebo run in to exclude placebo responders and those that did not, and for studies that excluded resistant patients and those that did not. The summary effect sizes

calculated in the respective subgroups were then compared in a Wald type test to assess if there were significant differences.

Meta-regressions were performed to determine whether year of publication, magnitude of standardized mean difference, dose, length and sample size of study, baseline severity, age or gender were associated with CVR. We also assessed collinearity between these potential moderating factors using correlation coefficients weighted by study size. Moderators that were found to be significant in the bivariate meta-regressions were then put into a single meta-regression model to see if they remained significant when assessed in a multiple variable model. The results of a meta-regression only allow one to see how the moderating factor is related to CVR, which is dependent on variability in both placebo and antidepressant arms. Therefore, in order to also examine whether the moderators primarily acted on antidepressant or placebo arms, we investigated the relationship between moderator and mean standardized variability in each placebo and active treatment arms separately (see further details regarding this variability measure in eMethods in Supplement):

$$Single\ arm\ standardised\ variability = ln(\delta_a/\bar{x}_a) = ln(S_a/\bar{x}_a) + \frac{1}{2(n_a - 1)}$$

Publication bias was assessed by visual inspection of funnel plots and the use of Egger's regression test. (Peters et al., 2006) I^2 values were calculated to quantify between-study inconsistency.

3. Results

3.1. Search strategy and studies selection

The search identified 9,389 articles. 134 studies met criteria for the meta-analysis (Figure 2). This included 17 unpublished studies, many involving reboxetine, a drug which the published evidence for has previously been shown to suffer from a potential risk of significant publication bias.(Eyding et al., 2010) The meta-analysis included data on 30,644 patients treated with antidepressants and 16,002 treated with placebo. There was a total of 227 separate antidepressant treatment arms and 122 placebo arms. The average age of subjects was 40.9 years, and males constituted 33.5% of trial participants. All the studies we included imposed a minimum baseline severity criterion, moderate severity (16~25 in converted HDRS-17 rating scale) in most cases.

3.2. Variability Difference between Antidepressant and Placebo

We found significantly lower variability in the antidepressant treated groups compared to placebo treated groups (CVR = 0.94, 95% CI 0.93-0.95, p<.0001, Figure 3). We next compared CVR between individual antidepressants. Imipramine showed significantly lower variability compared to fifteen other antidepressants. Moclobemide, amitriptyline and mirtazapine also showed significantly lower variability of change of symptom severity compared to several other antidepressants (Figure 3).

There was a positive relationship between change in total symptom and standard deviation of change (weighted r_p = 0.77, p<.0001, eFigure 1), indicating mean-scaling of variability. However, CVR, our primary outcome measure, adjusts for mean-scaling.(Nakagawa et al., 2015)

When not accounting for mean scaling, there was no significant difference in variability between antidepressant treated groups and placebo treated group (VR = 1.00, 95% CI 0.99-1.01, p=0.96; eFigure 2). Egger's test suggested the possibility of publication bias (z=-2.59, p=.01). A trim and fill analysis suggested no missing studies (eFigure 3A). I^2 was 0.00% indicating low levels of between-study inconsistency.

3.3. Efficacy Differences between Antidepressant and Placebo (SMD)

Antidepressants showed significantly greater improvements in symptoms than placebo (*g*=0.28, 95%CI 0.25-0.30, *p*<.0001) (Figure 4). Tests indicated moderate between study inconsistency (I^2=37.64% Q=369.1 *p*<.0001). Egger's test suggested the possibility of publication bias (*z*=3.54, *p*=.0004). The trim and fill analysis suggested 4 missing studies (eFigure 3B). However, including these putatively missing studies did not markedly change the results (*g*=0.27, 95%CI 0.24-0.30, *p*<.0001).

3.4. Subgroup analyses

When examining study quality, only 11 studies were classified as low risk in all domains of the Cochrane risk of bias tool. A subgroup analysis examining only these studies showed significantly reduced variability in change of symptom severity in the antidepressant relative to placebo group, consistent with our findings in the total dataset (CVR = 0.95, 95% CI 0.92-0.99, *p*<.001).

We investigated whether publication status affected CVR. Both published (CVR=0.93, 95%CI 0.92-0.94, *p*<.0001), and unpublished studies (CVR=0.97, 95%CI 0.94-0.99, *p*=.02)

showed significantly lower variability in symptom change in the antidepressant arm relative to the placebo arm.

We examined whether excluding placebo responders had an effect on CVR. Studies in which placebo responders were excluded (please see eTable 1 in Supplement), showed significantly lower variability in the antidepressant arm (CVR = 0.94, 95%CI 0.92-0.96, p<.0001), as did studies that did not exclude placebo responders (CVR = 0.93, 95%CI 0.92-0.95, p<.0001). However, there was no statistically significant difference in CVR between studies that did and did not exclude placebo responders (z=0.61, p=.54).

We also investigated whether excluding individuals who were potentially treatment-resistant affected CVR. Studies that specified treatment resistance as an exclusion criterion (eTable 1 of Supplement) showed significantly lower variability in the antidepressant arm relative to the placebo arm (CVR=0.94, 95%CI 0.92-0.96, p<.0001), and this was also the case in studies that did not have treatment-resistance as an exclusion criterion (CVR=0.94, 95%CI 0.92-0.95, p<.0001). There was no statistically significant difference between these two sets of studies (z=0.64, p=.52).

We next compared CVR between individual antidepressants. Imipramine showed significantly lower variability compared to fifteen other antidepressants. Moclobemide, amitriptyline and mirtazapine also showed significantly lower variability of change of symptom severity compared to several other antidepressants (Figure 5). We also compared CVR between antidepressants with multiple neurotransmitters profile and antidepressants with a primarily single neurotransmitter profile. We found that there was no difference in CVR between multiple neurotransmitter profile and single neurotransmitter profile antidepressants (z=-0.01, p=.99). For vortioxetine and nefazodone, although the NbN classifies them as predominantly acting on a single neurotransmitter system, they are sometimes classified as acting at multiple neurotransmitter systems.(Gonda et al., 2019; Protti et al., 2020) In view of this, we performed a subgroup analysis with vortioxetine and nefazodone in the multiple neurotransmitter profile group, and found that there was no difference in CVR between multiple neurotransmitter profile and single neurotransmitter profile antidepressants (z=-0.01, p=.99). Furthermore, we performed an analysis including only SSRIs as single neurotransmitter profile antidepressants, and found again that there was no difference in CVR between antidepressants with single/multiple neurotransmitter profile (z=-0.34, p=0.73).

3.5. Influence of study and patient characteristics on variability of antidepressant and placebo response

We found that there was a negative relationship between the standardized mean difference (SMD) calculated for a trial and its CVR (z=-7.21, p<.0001, Figure 6A). This indicates that trials demonstrating greater antidepressant efficacy were associated with lower variability in the antidepressant relative to placebo group. Further analysis of the individual trial arms showed placebo variability was not significantly associated with SMD (z=-0.84, p=.40, Figure 6B) but there was a significant relationship between greater SMD and reduced antidepressant variability (z=-6.48, p<.0001, Figure 6B).

There was a significant positive association between year of publication with CVR (z=3.01, p=.003) (Figure 6C), indicating older studies show less variability in symptom change in patients treated with antidepressants relative to placebo. Interestingly, variability in symptom change in both antidepressant (z=-2.97, p=.003) and placebo (z=-5.43, p<.0001) arms decreased with year of publication, but this has occurred at a faster rate in placebo arms (Figure 6D).

The meta-regression for baseline illness severity of patients included in studies showed no significant relationship with variability of change in symptom severity (z=1.23 p=.22) (Figure 6E). However, variability of change in symptom severity in both antidepressant (z=-2.59 p=.01) and placebo (z=-2.26 p=.02) arms decreased with increasing baseline severity. This indicates that studies included more severe patients showed lower variability of change in symptom severity in both antidepressant and placebo arms (Figure 6F).

The meta-regression for age indicated that there was no significant association between age and CVR (z=-1.40 p=.16) (Figure 6G). When examined separately according to treatment arms, variability in symptom change significantly decreased in the antidepressant arm with increasing age of patients (z=-3.49 p=.0005), but not in the placebo arm (z=-1.44 p=.15). This indicates that studies with older patients showed lower variability of symptom change in the antidepressant arms (Figure 6H).

Study duration (z=-0.01, p=.99), dose (z=-1.13, p=.26), and gender (male%) (z=-0.16, p=.87) did not show significant relationships with CVR.

A number of the variables found to be significant moderating factors in the meta-regressions showed a degree of collinearity. Larger sample sizes, greater illness severity and lower efficacy of antidepressant treatment are associated with more recent studies. We also found that more male patients and larger sample sizes are associated with lower efficacy (eTable 2). In a meta-regression of CVR including all the statistically significant moderating factors, SMD (z=-6.01, p<.0001) remained significant, but year of publication (z=-0.78, p=.44), baseline severity (z=0.69, p=.49), sample size of trial (z=0.19 p=.85), and gender (male%) (z=-0.64, p=.49) were no longer significant.

4. Discussion

We found that antidepressant treatment for major depressive disorder was associated with both significantly greater improvement (SMD), and also a more homogeneous change of symptom severity compared to placebo (CVR). Our finding regarding efficacy is consistent with previous meta-analyses,(Cipriani et al., 2018; Turner et al., 2008) but extends these by showing for the first time that there is more homogenous change in symptom severity change with antidepressants compared to placebo. In addition, we

found evidence that, compared to placebo, antidepressant-related improvement is more uniform in studies where antidepressants showed greater efficacy, and in older studies.

Our findings are not consistent with the hypothesis that there is greater variability in the antidepressant treatment effects because there are treatment responsive and non-responsive subtypes of major depression, this is a of great relevance when considering the feasibility of personalized medicine approaches.(Senn, 2018) One would expect that if subgroup phenomena were present this would be more apparent in studies that did not specifically exclude treatment resistant patients. In keeping with our main findings, however, this was not the case, and even when the analysis was restricted to studies that had not excluded treatment resistant patients there was no evidence of greater variability in the drug treated arm. Furthermore, we show that lower variability in symptom change is seen across all licensed antidepressants relative to placebo. Our findings show that there is no significant difference in variability of change of symptom severity between antidepressants with a multiple neurotransmitter profile and those with a single neurotransmitter profile, contrary to our hypothesis that multiple

neurotransmitter profile antidepressants are associated with more uniform change of symptom severity in patient samples. However, it should be noted that these comparisons were based on indirect comparisons. Head-to-head comparisons of antidepressants are needed to clarify this.

Clinical evidence indicates that the magnitude of placebo response in antidepressant trials has increased with time. (Rief et al., 2009; Weimer et al., 2015) Our findings showed that variability in symptom improvement in both antidepressant and placebo arms decreased in recent years, but this has occurred at a faster rate in placebo arms, our study extends recent research on the changing magnitude of placebo response (Weimer Katjia et al., 2015) to indicate that trials are increasingly including individuals who show a relatively homogenous, as well as high, placebo response.

Previous meta-analyses examining efficacy have found that efficacy is greater in studies including more severely ill patients.(Fournier et al., 2010b; Khan et al., 2002; Kirsch et al., 1999) However, we found that, relative to placebo, antidepressant variability was not significantly associated with baseline severity. Furthermore, variability in symptom improvement in both antidepressant and placebo arms was lower in studies including

patients with more severe symptoms at baseline. One possible explanation for this finding is that with increasing levels of severity a large placebo response becomes less common, and our finding is driven by lower variability in the placebo arm. This explanation is supported by studies showing an inverse relationship between placebo response and depression severity, with less placebo response in more severely depressed patients.(Khan et al., 2002; Kirsch et al., 1999)

Despite contacting authors to obtain missing data, measures of variance were not available for some RCTs. Variance data were also missing for subscales of the rating scales, and as such we could not examine different symptom domains. All of the studies we included imposed a minimum baseline severity criterion (moderate severity in most cases), so our results may not generalize to milder forms of depression. A further limitation is that the mean duration of trials we included was 8 weeks, so it is unclear how our findings apply to long-term use of antidepressants. We found evidence of publications bias in some analyses but, as variability was not an outcome of interest in the trials, it is unlikely that this was a bias against publishing studies that show greater variability with antidepressant response. However, due to the association between CVR and SMD, some bias could still be present. Nevertheless, our subgroup analyses

indicated that findings would not change substantially with the inclusion of putatively missing studies. It is important to note that, although there was less variability in the antidepressant treatment group, there are, nevertheless, some patients who show minimal response. A key question for future studies is whether this represents random variation in response, or if there are patients who consistently show minimal response to treatment, including placebo. Repeated period cross-over trials are needed to address this question.

The possibility exists that the lower variability of change of symptom severity observed in the antidepressant arm is an artefact of trial inclusion criteria which might be designed with the aim of recruiting individuals likely to respond to antidepressants. However, this was not supported by the subgroup analyses excluding trials that specifically excluded treatment resistant patients. Also, the lower variability in active treatment arms observed in the context of RCTs may not accurately reflect the variability of treatment in real-world clinical settings, and our findings cannot exclude the possibility that there is a subtype of depression that does not enter clinical trials. This is a potential limitation of clinical trial design generally. Pragmatic trial designs that recruit more representative patients are one strategy to address this limitation. To fully

address the question of subtypes, future work will require markedly different trial designs such as repeated N-of-1 studies.(Senn, 2018) It should also be recognized that our findings are specific to treatment response, and so do not exclude the existence of sub-types of depression linked to factors other than treatment response.

In summary, we found a more homogeneous change of symptom severity in those receiving antidepressants relative to those receiving placebo, as well as greater symptomatic improvement with antidepressant treatment. These findings indicate a more uniform effect of antidepressant treatment that exceeds non-specific effects seen with placebo treatment.

Disclosures

XG, RM, TP, SN, and KB declare no financial conflicts of interest. Dr. Mizuno has received manuscript fees or speaker's honoraria from Sumitomo Dainippon Pharma and Yoshitomi Yakuhin, fellowship grants from Japan Society for the Promotion of Science, Astellas Foundation for Research on Metabolic Disorders, Japanese Society of Clinical Neuropsychopharmacology, and Mochida Memorial Foundation for Medical and

Pharmaceutical Research, and consultant fees from Bracket within the past three years. OH has received investigator-initiated research funding from and/or participated in advisory/ speaker meetings organised by Astra-Zeneca, Autifony, BMS, Eli Lilly, Heptares, Jansenn, Lundbeck, Lyden-Delta, Otsuka, Servier, Sunovion, Rand and Roche. Neither Dr Howes or his family have been employed by or have holdings/ a financial stake in any biomedical company.

Acknowledgments

XG's work is funded by the National Key R&D Program of China (no. 2018YFC1314600), and fellowship grants from Chinese Scholarship Council. OH's work is supported by Medical Research Council-UK (no. MC-A656-5QD30), Maudsley Charity (no. 666), Brain and Behavior Research Foundation, and Wellcome Trust (no. 094849/Z/10/Z) grants to Dr Howes and the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. RM's research is funded by the Wellcome Trust (no. 200102/Z/15/Z) and a NIHR clinical lectureship.

References

Bollini, P., Pampaliona, S., Tibaldi, G., Kupelnick, B., Munizza, C., 1999. Effectiveness of antidepressants: meta-analysis of dose-effect relationships in randomised clinical trials. The British Journal of Psychiatry 174, 297-303.

Brugger, S.P., Howes, O.D., 2017. Heterogeneity and Homogeneity of Regional Brain Structure in Schizophrenia: A Meta-analysis. JAMA psychiatry.

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., Saxena, S., 2016. Scaling-up treatment of depression and anxiety: a global return on investment analysis. The Lancet Psychiatry 3, 415-424.

Cipriani, A., Furukawa, T.A., Salanti, G., Chaimani, A., Atkinson, L.Z., Ogawa, Y., Leucht, S., Ruhe, H.G., Turner, E.H., Higgins, J.P.T., Egger, M., Takeshima, N., Hayasaka, Y., Imai, H., Shinohara, K., Tajika, A., Ioannidis, J.P.A., Geddes, J.R., 2018. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet.

Conway, C.R., George, M.S., Sackeim, H.A., 2017. Toward an Evidence-Based, Operational Definition of Treatment-Resistant Depression: When Enough Is Enough. JAMA psychiatry.

Eyding, D., Lelgemann M Fau - Grouven, U., Grouven U Fau - Härter, M., Härter M Fau - Kromp, M., Kromp M Fau - Kaiser, T., Kaiser T Fau - Kerekes, M.F., Kerekes Mf Fau - Gerken, M., Gerken M Fau - Wieseler, B., Wieseler, B., 2010. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. BMJ.

Fournier, J.C., DeRubeis Rj Fau - Hollon, S.D., Hollon Sd Fau - Dimidjian, S., Dimidjian S Fau - Amsterdam, J.D., Amsterdam Jd Fau - Shelton, R.C., Shelton Rc Fau - Fawcett,

J., Fawcett, J., 2010a. Antidepressant drug effects and depression severity: a patient-level meta-analysis. JAMA.

Fournier, J.C., DeRubeis, R.J., Hollon, S.D., Dimidjian, S., Amsterdam, J.D., Shelton, R.C., Fawcett, J., 2010b. Antidepressant drug effects and depression severity: a patient-level meta-analysis. Jama 303, 47-53.

Friedrich, M., 2017. Depression is the leading cause of disability around the world. Jama 317, 1517-1517.

Gonda, X., Sharma, S.R., Tarazi, F.I., 2019. Vortioxetine: a novel antidepressant for the treatment of major depressive disorder. Expert Opin Drug Discov.

Hayasaka, Y., Purgato, M., Magni, L.R., Ogawa, Y., Takeshima, N., Cipriani, A., Barbui, C., Leucht, S., Furukawa, T.A., 2015. Dose equivalents of antidepressants: Evidence-based recommendations from randomized controlled trials. Journal of affective disorders 180, 179-184.

Higgins, J.P., Altman, D.G., 2008. Assessing risk of bias in included studies. Cochrane handbook for systematic reviews of interventions: Cochrane book series, 187-241.

Higgins, J.P., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., Savović, J., Schulz, K.F., Weeks, L., Sterne, J.A., 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Bmj 343, d5928.

Jakubovski, E., Varigonda, A.L., Freemantle, N., Taylor, M.J., Bloch, M.H., 2015. Systematic review and meta-analysis: dose-response relationship of selective serotonin reuptake inhibitors in major depressive disorder. American Journal of Psychiatry 173, 174-183.

Khan, A., Leventhal Rm Fau - Khan, S.R., Khan Sr Fau - Brown, W.A., Brown, W.A., 2002. Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. Journal of clinical psychopharmacology.

Kirsch, I., Sapirstein, G., 1999. Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medications.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS medicine 6, e1000100.

McCutcheon, R.A.-O., Pillinger, T., Mizuno, Y.A.-O., Montgomery, A., Pandian, H., Vano, L.A.-O.X., Marques, T.R., Howes, O.D., 2019. The efficacy and heterogeneity of antipsychotic response in schizophrenia: A meta-analysis. LID - 10.1038/s41380-019-0502-5 [doi]. Mol Psychiatry.

Montgomery, S.A., Åsberg, M., 1979. A new depression scale designed to be sensitive to change. The British journal of psychiatry 134, 382-389.

Nakagawa, S., Poulin, R., Mengersen, K., Reinhold, K., Engqvist, L., Lagisz, M., Senior, A.M., 2015. Meta - analysis of variation: ecological and evolutionary applications and beyond. Methods in Ecology and Evolution 6, 143-152.

Otte, C., Gold, S.M., Penninx, B.W., Pariante, C.M., Etkin, A., Fava, M., Mohr, D.C., Schatzberg, A.F., 2016. Major depressive disorder. Nature reviews. Disease primers.

Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R., Rushton, L., 2006. Comparison of two methods to detect publication bias in meta-analysis. Jama 295, 676-680.

Protti, M., Mandrioli, R.A.-O.X., Marasca, C., Cavalli, A., Serretti, A., Mercolini, L., 2020. New-generation, non-SSRI antidepressants: Drug-drug interactions and therapeutic drug monitoring. Part 2: NaSSAs, NRIs, SNDRIs, MASSAs, NDRIs, and others. LID - 10.1002/med.21671 [doi]. Med Res Rev.

Rief, W., Nestoriuc Y Fau - Weiss, S., Weiss S Fau - Welzel, E., Welzel E Fau - Barsky, A.J., Barsky Aj Fau - Hofmann, S.G., Hofmann, S.G., 2009. Meta-analysis of the placebo response in antidepressant trials. Journal of affective disorders.

Rush, A.J., Trivedi Mh Fau - Wisniewski, S.R., Wisniewski Sr Fau - Nierenberg, A.A., Nierenberg Aa Fau - Stewart, J.W., Stewart Jw Fau - Warden, D., Warden D Fau - Niederehe, G., Niederehe G Fau - Thase, M.E., Thase Me Fau - Lavori, P.W., Lavori Pw Fau - Lebowitz, B.D., Lebowitz Bd Fau - McGrath, P.J., McGrath Pj Fau - Rosenbaum, J.F., Rosenbaum Jf Fau - Sackeim, H.A., Sackeim Ha Fau - Kupfer, D.J., Kupfer Dj Fau - Luther, J., Luther J Fau - Fava, M., Fava, M., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. American Journal of Psychiatry.

Senn, S., 2018. Statistical pitfalls of personalized medicine. Nature Publishing Group.

Team, R.C., 2013. R: A language and environment for statistical computing.

Turner, E.H., Matthews, A.M., Linardatos, E., Tell, R.A., Rosenthal, R., 2008. Selective publication of antidepressant trials and its influence on apparent efficacy. New England Journal of Medicine 358, 252-260.

Uchida, H., Yamawaki, S., Bahk, W.M., Jon, D.I., 2016. Neuroscience-based Nomenclature (NbN) for Clinical Psychopharmacology and Neuroscience. Clin Psychopharmacol Neurosci.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. Journal of statistical software 36, 1-48.

Weimer, K., Colloca, L., Enck, P., 2015. Placebo effects in psychiatry: mediators and moderators. The Lancet Psychiatry 2, 246-257.

Weimer Katjia, L., Colloca, E.P., 2015. Placebo eff ects in psychiatry: mediators and moderators. Lancet Psychiatry.

WHO, 2017. Depression and other common mental disorders: global health estimates. World Health Organization.

Williams, J.B., 1988. A structured interview guide for the Hamilton Depression Rating Scale. Archives of general psychiatry 45, 742-747.

Willner, P., Scheel-Krüger J Fau - Belzung, C., Belzung, C., 2013. The neurobiology of depression and antidepressant action. Neurosci Biobehav Rev.

Winkelbeiner, S., Leucht, S., Kane, J.M., Homan, P., 2019. Evaluation of Differences in Individual Treatment Response in Schizophrenia Spectrum Disorders: A Meta-analysis. JAMA psychiatry.

Figure titles & legends

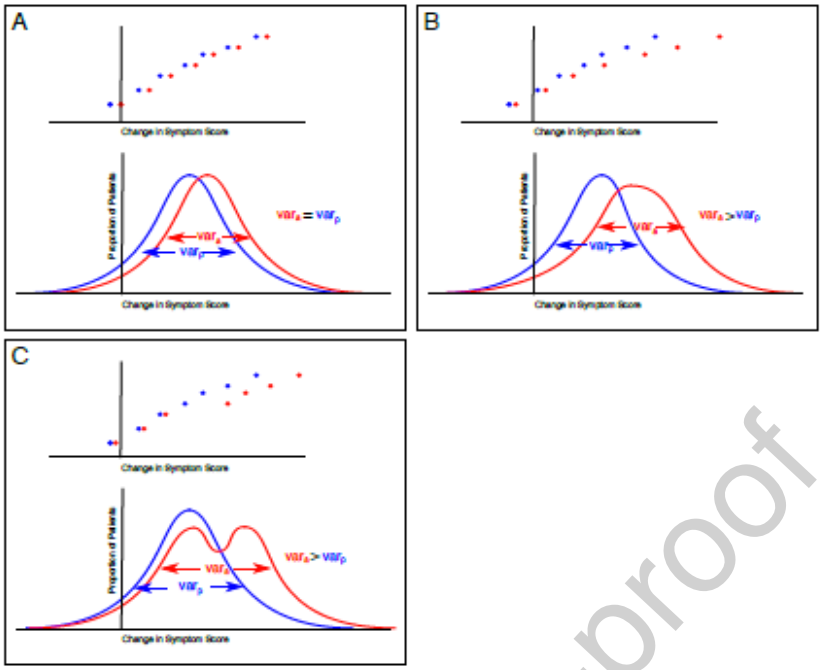

Figure 1. Illustration of different potential models of placebo and antidepressant response. Red illustrates patients receiving active treatment and blue illustrates patients receiving placebo treatment. Var_a and Var_p represent variability of antidepressant and placebo treatment respectively. The upper graphs plot individual subjects receiving placebo/active treatment. The lower graphs illustrate the corresponding group distributions. In Panel A, all subjects have the same small 'drug response', whereas in Panel B, some individuals show a much greater drug response than others. Both scenarios have similar mean differences between placebo and active treatment, but the variance is greater for drug treatment in the second scenario. There may also exist a sub-population of patients that shows no response beyond that seen with a placebo, resulting in a bimodal distribution and increased variance in response in the active relative to the placebo group (Panel C).

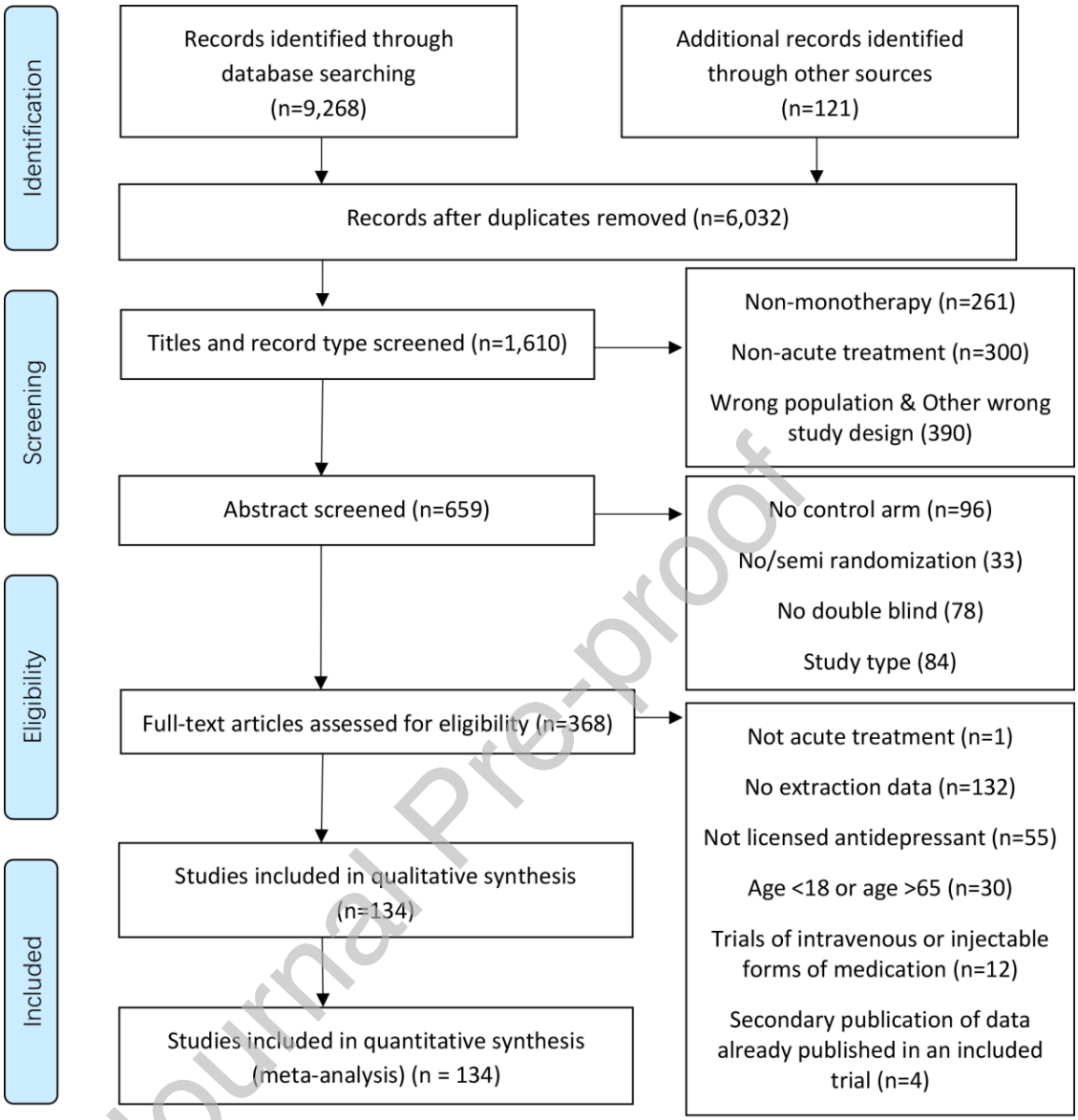

Figure 2. Coefficient of variation (CVR) for change in symptom for placebo-controlled trials of antidepressant treatment. CVR was significantly lower in antidepressant treated relative to placebo treated patients (CVR = 0.94, 95% CI 0.93-0.95, p<.0001), indicating lower variability in symptom change with antidepressant treatment compared to the placebo. Abbreviations: CI, confidence interval; CVR, coefficient of variation

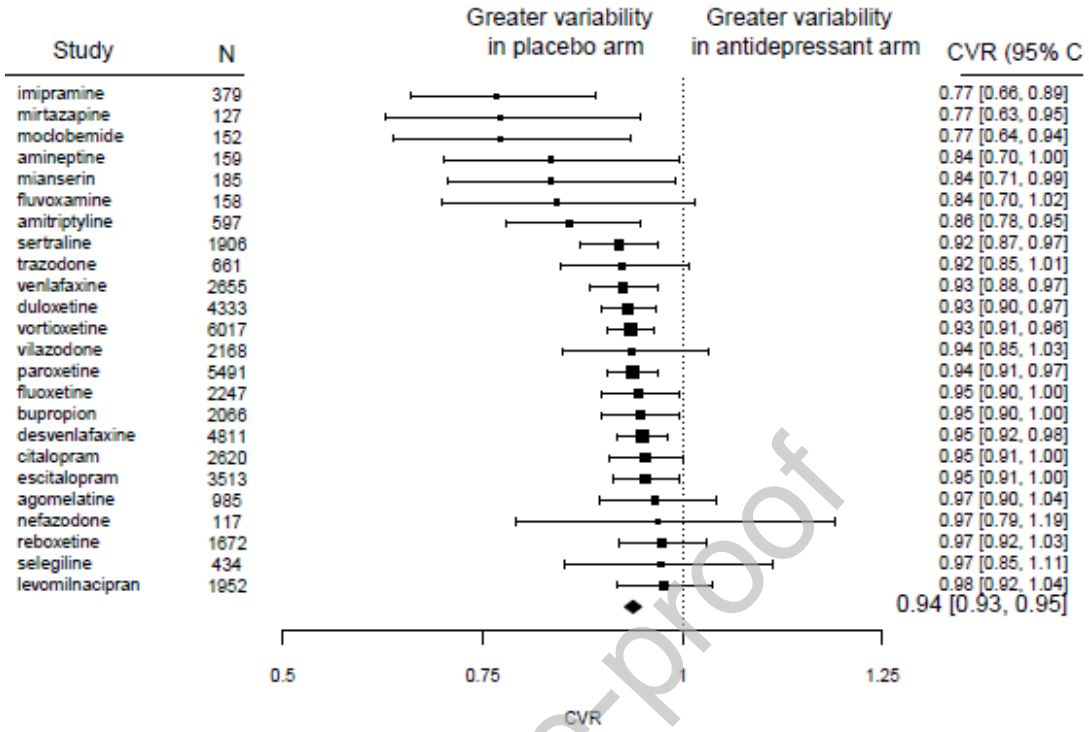

Figure 3. Efficacy difference (SMD) between antidepressant and placebo.

Antidepressants showed greater improvements in symptoms than placebo (g=0.28, 95%CI 0.25-0.30, p<.0001). Abbreviations: CI, confidence interval; SMD, standardized mean difference

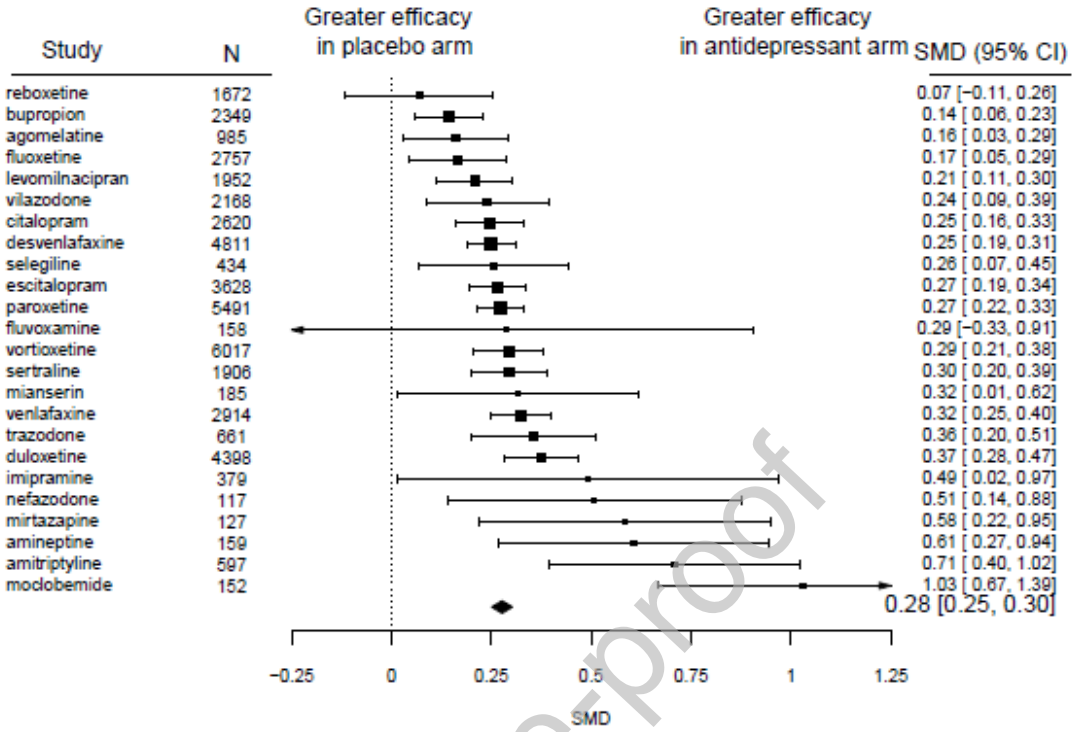

Figure 4. Comparison of the coefficient of variation ratio (CVR) between different individual antidepressant. Asterisks indicate whether the comparison is statistically significant at p<.05. Red indicates that the variability for the antidepressant named on the y axis is lower than the antidepressant named on the x axis, blue indicates that the variability of the antidepressant named on the x axis is lower than the antidepressant named on the y axis. Imipramine, moclobemide, amitriptyline and mirtazapine show significantly lower variability of symptom change than a number of other individual antidepressant, with statistical significance ranging from (z=3.28, p=.001) for bupropion

compared with imipramine, to (z=-1.99, p= .0463) for desvenlafaxine compared with amitriptyline.

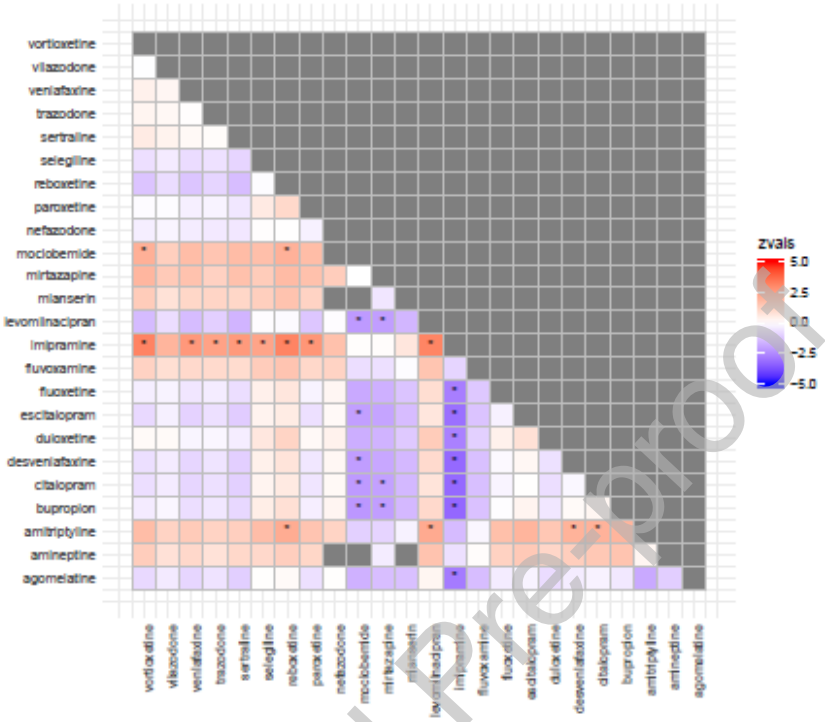

Figure 5. **Meta-regression exploring Influence of study and patient characteristics on variability of antidepressant and placebo response.** Figures on the left show meta-regressions exploring the influence of trial characteristics on the coefficient of variation (CVR) for change in total symptoms. Figures on the right study the same trial characteristics but examine placebo and antidepressant arms separately. (A) Trials that show greater antidepressant efficacy (greater mean symptom improvement with antidepressant relative to placebo) show lower variability of symptom change in antidepressant treatment relative to placebo (z=-7.21, p<.0001). (B) Trials showing

greater efficacy show reduced symptom change variability in antidepressant treated arms (z=-6.47, p<.0001), but no significant relationship is apparent with regard to placebo variability (z=-0.84, p=.40). (C) Figure shows that there is a positive relationship between year of publication and variability of symptom change in patients receiving antidepressant treatment relative to placebo, indicating older trials show lower variability in symptom change with (z=3.01, p=.003). (D) Showing that there is a negative relationship between year of publication and symptom change variability in symptom change in both antidepressant (z=-2.97, p=.003) and placebo (z=-5.43, p<.0001) arms. (E) No significant association is found between baseline illness severity and symptom change variability in antidepressant arms relative to placebo arms (z=-1.23, p=.22). F) Both antidepressant (z=-2.59 p=.01) and placebo (z=-2.26 p=.02) arms show decreasing variability with increasing baseline severity, indicating in studies included more severe patients showed lower variability of symptom change in both antidepressant and placebo arms. (G) Figure shows no significant association between age and CVR (z=-1.40 p=.16). (H) Showing that variability in symptom change with antidepressants decrease with increasing age (z=-3.49 p=.0005), at a faster rate than placebo variability (z=-1.44 p=0.15). Abbreviations: SMD, standardized mean difference

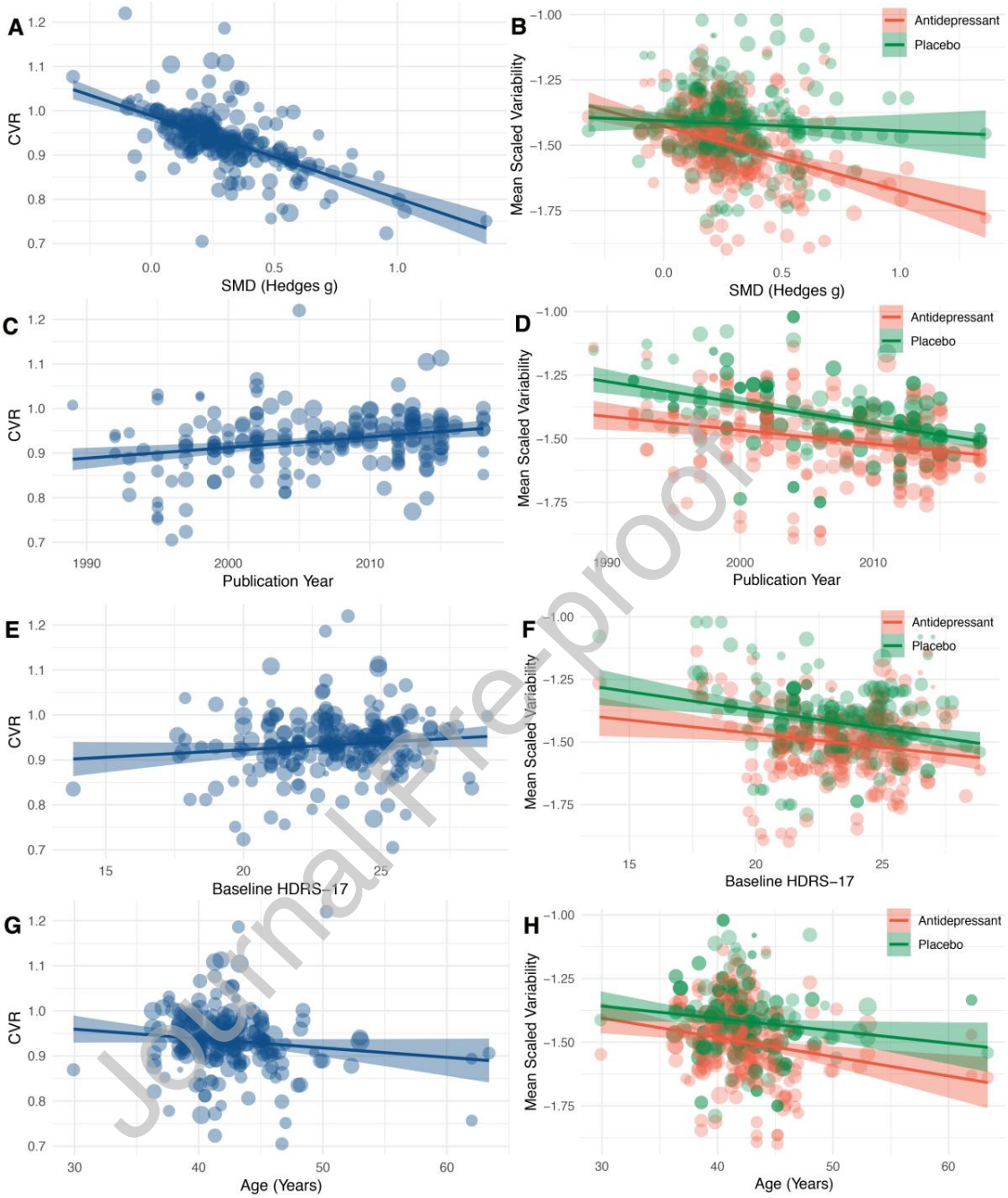

Figure 6