JAMA Psychiatry | Original Investigation

Individual Differences in Response to Antidepressants
A Meta-analysis of Placebo-Controlled Randomized Clinical Trials

Marta M. Maslej, PhD; Toshiaki A. Furukawa, MD, PhD; Andrea Cipriani, MD, PhD; Paul W. Andrews, PhD, JD; Marcos Sanches, MSc; Anneka Tomlinson, PhD, MD; Constantin Volkmann, MD; Robert A. McCutcheon, PhD; Oliver Howes, PhD, DM; Xin Guo, PhD; Benoit H. Mulsant, MD, MS

IMPORTANCE Antidepressants are commonly used to treat major depressive disorder (MDD). Antidepressant outcomes can vary based on individual differences; however, it is unclear whether specific factors determine this variability or whether it is at random.

OBJECTIVE To investigate the assumption of systematic variability in symptomatic response to antidepressants and to assess whether variability is associated with MDD severity, antidepressant class, or study publication year.

DATA SOURCES Data used were updated from a network meta-analysis of treatment with licensed antidepressants in adults with MDD. The Cochrane Central Register of Controlled Trials, CINAHL, Embase, LILACS database, MEDLINE, MEDLINE In-Process, and PsycInfo were searched from inception to March 21, 2019. Additional sources were international trial registries and sponsors, drug companies and regulatory agencies' websites, and reference lists of published articles. Data were analyzed between June 8, 2020, and June 13, 2020.

STUDY SELECTION Analysis was restricted to double-blind, randomized placebo-controlled trials with depression scores available at the study's end point.

DATA EXTRACTION AND SYNTHESIS Baseline means, number of participants, end point means and SDs of total depression scores, antidepressant type, and publication year were extracted.

MAIN OUTCOMES AND MEASURES Log SDs ($b_{ln}\,\hat{\sigma}$) were derived for treatment groups (ie, antidepressant and placebo). A random-slope mixed-effects model was conducted to estimate the difference in $b_{ln}\,\hat{\sigma}$ between treatment groups while controlling for end point mean. Secondary models determined whether differences in variability between groups were associated with baseline MDD severity; antidepressant class (selective serotonin reuptake inhibitors and other related drugs; serotonin and norepinephrine reuptake inhibitors; norepinephrine-dopamine reuptake inhibitors; noradrenergic agents; or other antidepressants); and publication year.

RESULTS In the 91 eligible trials (18 965 participants), variability in response did not differ significantly between antidepressants and placebo ($b_{ln}\,\hat{\sigma}$, 1.02; 95% CI, 0.99-1.05; P = .19). This finding is consistent with a range of treatment effect SDs (up to 16.10), depending on the association between the antidepressant and placebo effects. Variability was not associated with baseline MDD severity or publication year. Responses to noradrenergic agents were 11% more variable than responses to selective serotonin reuptake inhibitors ($b_{ln}\,\hat{\sigma}$, 1.11; 95% CI, 1.01-1.21; P = .02).

CONCLUSIONS AND RELEVANCE Although this study cannot rule out the possibility of treatment effect heterogeneity, it does not provide empirical support for personalizing antidepressant treatment based solely on total depression scores. Future studies should explore whether individual symptom scores or biomarkers are associated with variability in response to antidepressants.

Editorial page 478

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Benoit H. Mulsant, MD, MS, Centre for Addiction and Mental Health, University of Toronto, 250 College St, Ste 835, Toronto, ON M5T 1R8, Canada (benoit.mulsant@utoronto.ca).

JAMA Psychiatry. 2021;78(5):490-497. doi:10.1001/jamapsychiatry.2020.4564
Published online February 17, 2021.

Major depressive disorder (MDD) is a common and heterogeneous mental condition characterized by emotional, cognitive, somatic, and behavioral symptoms.[1] Antidepressants are one of the first-line interventions for the treatment of depression,[2] but their efficacy appears to be variable. While a significant proportion of individuals experience remission of depression after 8 weeks of treatment, more than 50% of patients improve very little or their depression worsens.[3] This observed variation has prompted efforts to identify moderators of antidepressant efficacy and to personalize treatments by matching specific antidepressants with the unique characteristics of individual patients.[4-6] The variability in the efficacy of psychiatric medications is typically deduced from aggregate data of randomized clinical trials (RCTs), which estimate average treatment effects.[5,6] To some degree, treatment effects in RCTs vary between individuals owing to random factors or other factors, such as placebo effects, regression to the mean, or measurement error.[6,7] The ability to personalize treatments rests on the assumption that individual differences contribute to this variability, but detecting treatment by individual interactions requires more complex study designs.[8]

Despite the paucity of studies designed to detect treatment by individual interactions, there is a widely held assumption that individual differences moderate the effect of antidepressants on depressive symptoms (ie, response).[9-12] Data from RCTs show that individuals with symptoms of depression assigned to receive the same antidepressant at the same dose and for the same period can experience very different outcomes.[5] The source of this variability is believed to result from individual differences in clinical or biological factors.[10] If indeed some antidepressants (or classes of antidepressants) are more effective at treating patients whose MDD is characterized by specific clinical or biological factors, such differences would be consistent with the assumption that individuals vary systematically in their response to antidepressants. However, if this variability is driven by other factors (eg, placebo response and measurement error), it may not be possible to personalize antidepressant treatment. Given that the potential for personalization rests on the validity of an assumption of systematic variability in observed response to antidepressants, it is important to evaluate it rigorously. Thus, we compared variability in observed outcomes of patients with MDD assigned to receive antidepressants or placebo to assess whether the observed variability in response to antidepressants is owing to systematic, nonrandom factors. We hypothesized that if variability in observed response to antidepressants includes an individual by treatment interaction, it would differ from the variability in observed response to placebo.

A vigorous methodologic debate has been taking place on the best way to assess variability in response to psychiatric treatment.[13-17] This debate involves how to deal with the problem that means and their variability (ie, SDs) are often not independent.[13,17,18] Several previously published articles[19,20] have relied on methods that make assumptions about the nature of the mean-SD relationship, which could lead to biased estimates if these assumptions are not met. To address this issue, we use a random-slope mixed-effects model (RSMM) that accounted for the association between end point depression scores and variability by modeling this association directly from the data.[13,18] Furthermore, we updated the open data set of RCTs evaluating outcomes of antidepressants in patients with MDD from Cipriani et al[21] that has been used in prior studies of variability in response to antidepressants.[13,19,20]

We also examined whether baseline severity of depression, antidepressant class, or the year in which studies were published is associated with variability in response to antidepressants. Because on average, the effects of antidepressants are modest,[21] we expected variability to increase for moderators associated with increased response. Thus, we examined whether responses in groups of participants whose symptoms were initially more severe would be more variable because the effects of antidepressants may be more pronounced in individuals with severe depression.[22,23] We also expected that variability might differ based on the way that different antidepressant classes interact with different neurotransmitter systems. Specifically, antidepressants affecting multiple neurotransmitter systems might produce more variable outcomes than would antidepressants with more selective effects.

Key Points

Question Is there evidence that response to antidepressants varies based on individual differences?

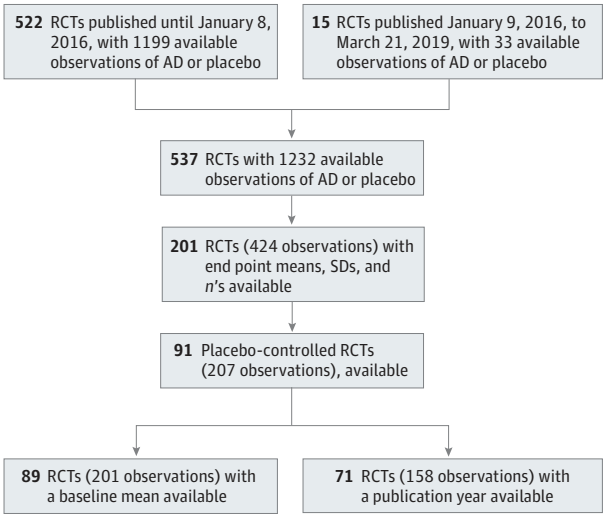
Findings In this meta-analysis of 91 randomized clinical trials (18 965 participants) on the use of antidepressants in major depression, no evidence was found of more variability in response to antidepressants than to placebo. Variability was not associated with baseline depression severity or study year, but variability in response to noradrenergic agents was higher than that of selective serotonin reuptake inhibitors.

Meaning Individual differences may not underlie variability in the association between total depression scores and antidepressant treatment; future efforts toward personalization should focus on individual symptoms or biomarkers.

Methods

We started with publicly available data from a published network meta-analysis of 522 RCTs evaluating the effects of antidepressants on MDD.[21] The methods and descriptive statistics for this meta-analysis are published elsewhere.[21,24] Briefly, selected databases (the Cochrane Central Register of Controlled Trials, CINAHL, Embase, LILACS database, MEDLINE, MEDLINE In-Process, and PsycInfo) were searched from their inception to January 8, 2016, using terms that included references to depression in combination with a list of antidepressants. Additional sources were international trial registries, drug approval agency websites, and key scientific journals. We updated the search to include 15 additional RCTs published between January 9, 2016, and March 21, 2019, resulting in a total of 537 RCTs (**Figure 1**). Of these 537 RCTs, 256 (48%) were conducted in North America, 142 (26%) in Europe, and 42 (8%) in

Figure 1. Study Selection Process



AD indicates antidepressant; RCT, randomized clinical trial.

Asia, with the remaining studies being cross-continental or from other regions. There was a total of 89 838 participants allocated to an antidepressant and 30 251 allocated to a placebo. Participants' mean (SD) age was 44.18 (11.21) years, and 62% were women. The included studies assessed depressive symptoms using one of several versions of the Hamilton Rating Scale for Depression (HAMD-17,[25] HAMD-21,[25] HAMD-24,[26] HAMD-29,[27] and HAMD-31[28]), the *Montgomery Åsberg* Depression Rating Scale,[29] or the Inventory of Depressive Symptomatology.[30] End point scores were extracted as close to 8 weeks after the start of antidepressant treatment or placebo as possible, with the median duration of extracted scores being 8 weeks (interquartile range, 6-8 weeks). The data were analyzed between June 8, 2020, and June 13, 2020.

Eligibility Criteria

Our analysis included observations from placebo-controlled RCTs with available data at end point (means and SDs of total depression scores and number of participants assessed in each group). Figure 1 depicts our selection process and the resulting number of included RCTs. From the publicly available data[21] corresponding to the eligible RCTs, we extracted baseline and end point means, SDs, number of participants in each group, antidepressant drug, and, when available, the year of RCT publication.

Statistical Analysis

Primary Analysis

Based on investigations into the appropriate approach to quantify variability in our included RCTs (described in the eMethods in the Supplement), we used an RSMM to estimate differences in variability between treatment groups, while controlling for the end point mean score.[13,18] Following Nakagawa et al,[18] we used an unbiased estimator of the natural logarithm of the population SD and its sampling variance.[31] For each observation corresponding to an antidepressant or placebo group, we calculated the log SD:

$$\hat{\sigma} = \ln s + \frac{1}{2(n-1)}$$

where s refers to the SD and n refers to the number of participants for that group.[18]

We used the following formula to derive sampling variances:

$$s^2_{ln\,\hat{\sigma}} = \frac{1}{2(n-1)}$$

where n refers to the number of participants for that group.[18] In our primary RSMM, we specified each ln $\hat{\sigma}$ (fitted with its associated sampling variance) as the response variable. We specified treatment group as a categorical predictor indicating whether the ln $\hat{\sigma}$ corresponded to antidepressant (coded as 1) or placebo (coded as 0). To account for the mean-SD association, we added the log end point mean score (z-transformed across the entire data set) as a predictor. Because the effect of treatment group, estimated by $b_{ln}\,\hat{\sigma}$, represented the difference in ln $\hat{\sigma}$ between antidepressant and placebo,[18] we report the exponentiated $b_{ln}\,\hat{\sigma}$. A group effect larger than 1 indicated higher variability in antidepressant groups than placebo groups; conversely, a group effect lower than 1 indicated less variability in the antidepressant groups compared with placebo groups.[18] We used methods from previous work[7,13] to explore possible values for treatment effect SD, based on findings of this primary analysis and a range of possible associations between the treatment and placebo effects.

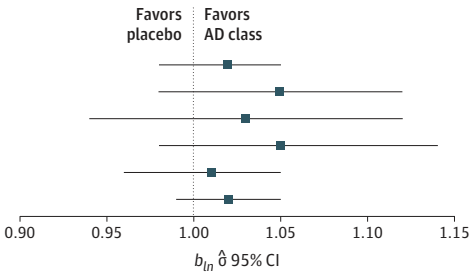
Secondary Analyses

We repeated our primary analyses adding a treatment group by baseline depression interaction. We used all observations with a baseline depression severity score available (Figure 1) and z-transformed log baseline means across the entire data set. A treatment group by baseline depression interaction effect larger than 1 was consistent with our hypothesis, indicating more variability in response to antidepressant as compared with placebo with increasing baseline depression severity.

For each available observation from the eligible RCTs, we categorized the antidepressants into 1 of the following 5 classes based on their main putative mechanisms of action[32,33]: selective serotonin reuptake inhibitors (SSRIs) and other related drugs (citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline, and vilazodone); serotonin and norepinephrine reuptake inhibitors (SNRIs) (desvenlafaxine and venlafaxine); norepinephrine-dopamine reuptake inhibitors (NDRIs) (bupropion); noradrenergic agents (NAs) (amitriptyline and reboxetine); and other antidepressants (agomelatine, mirtazapine, and trazodone). Randomized clinical trials reporting observations for 4 antidepressants (ie, duloxetine, levomilnacipran, nefazadone, and vortioxetine) did not report end point means and

Figure 2. Comparisons of Variability Between Antidepressants (ADs) and Placebo, Separated by Class

AD class	No. of RCTs	No. of observations available	$b_{ln}\hat{\sigma}$ (95% CI)	P value
SSRI	47	56	1.02 (0.98-1.05)	.28
SNRI	9	11	1.05 (0.98-1.12)	.11
NDRI	10	11	1.03 (0.94-1.12)	.55
NA	18	18	1.05 (0.98-1.14)	.22
Other AD	18	23	1.01 (0.96-1.05)	.80
Any AD	91	207	1.02 (0.99-1.05)	.19



Comparisons of placebo (the reference group) and AD (or AD class) indicate the difference between the 2 groups in *ln* $\hat{\sigma}$. These estimates have been exponentiated, with a $b_{ln}\hat{\sigma}$ less than 1 representing less variability than placebo. NA indicates norepinephrinergic agent; NDRI, norepinephrine-dopamine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; SNRI, serotonin and norepinephrine reuptake inhibitor.

SDs and were not included in our analysis. We repeated our primary analyses replacing the categorical treatment group predictor with antidepressant class. We specified placebo as the reference group, so that an effect of antidepressant class (exponentiated) larger than 1 indicated greater variability in symptomatic response for that class than placebo. To compare variability in response among antidepressant classes, we restricted our analyses to antidepressant groups only, specifying SSRIs as the reference group. Any antidepressant class effects larger than 1 were consistent with our hypothesis, indicating greater symptomatic variability for classes with less selective effects than SSRIs.

Finally, we investigated whether the year in which RCTs were published was associated with differences in variability. For RCTs with a publication year available (Figure 1), we rescaled years by subtracting the median year (ie, 2001). We conducted an RSMM including a year by treatment group interaction, with an interaction effect greater than 1 indicating more variability in response to antidepressant than placebo in RCTs conducted more recently.

Sensitivity Analyses

In the included RCTs, baseline and end point means were measured using the HAMD-17, HAMD-21, HAMD-24, and MADRS, introducing the possibility of spurious effects associated with the use of different scales.[13] The duration of treatment also differed between RCTs (ranging from 4-12 weeks). To examine whether different scales and treatment durations affected our results, we repeated our primary and secondary analyses adding scale as a categorical predictor and duration of treatment in weeks as a continuous predictor.

All analyses were completed in R, version 3.5.3 (the R Foundation),[34] using the MCMCglmm package.[35] We used functions provided by Nakagawa et al[18] to calculate ln $\hat{\sigma}$ and sampling variances and the *MCMCglmm* function for generating RSMMs.[35] Detailed descriptions of model specifications and priors are provided in the eMethods in the Supplement. To ensure that choice of priors did not affect our results, we repeated our primary and secondary analyses with expanded priors.[18] The significance threshold was .05, and significance testing was 2-sided. To ensure reproducibility, our code[36] is freely available online. The updated data set is available on request from the authors.

Results

Primary Analysis

A total of 91 RCTs comprising 18 965 unique participants met our inclusion criteria. To measure outcomes at end point, 44 RCTs (47%) used the HAMD-17 and 33 (37%) used the HAMD-21. Because some RCTs compared placebo with more than 1 antidepressant (ie, multiarm trials), 207 observations were available (Figure 1). End point means were associated with variability ($b_{ln}\hat{\sigma}$ = 1.07; 95% CI, 1.04-1.10; P < .001). We tested for nonlinearity in this association by adding a quadratic term for end point means, and there was no evidence that adding this term improved the model (eTable 1 in the Supplement). After controlling for the linear association, there was 2% more variability in responses to antidepressants compared with placebo, but this difference was not statistically significant ($b_{ln}\hat{\sigma}$ = 1.02; 95% CI, 0.99-1.05; P = .19) (**Figure 2**).

We calculated the upper bounds on the treatment effect SD, based on this finding and a range of possible associations between the treatment and placebo effects (details of this calculation and simulation are provided in the eResults and eFigures 3-5 in the Supplement). These analyses suggest that $b_{ln}\hat{\sigma}$ = 1.05 (the upper limit of our estimated $b_{ln}\hat{\sigma}$) is compatible with a range of treatment effect SDs (eFigure 3 in the Supplement) up to 16.10 (eTable 1 in the Supplement).

Secondary Analyses

Depression Severity

There were 89 placebo-controlled RCTs with both baseline and end point means available, corresponding to 201 observations (Figure 1). We found no evidence that the association between treatment group and variability was associated with baseline depression severity ($b_{ln}\hat{\sigma}$ = 0.99; 95% CI, 0.96, 1.02; P = .58) (results for model predictors are provided in the **Table**).

Antidepressant Class

Figure 2 provides the number of RCTs reporting outcomes for each antidepressant class and the number of available observations. Responses to all antidepressant classes were not more variable than responses to placebo (depicted in Figure 2). Responses to NDRIs, SNRIs, or other antidepressants were not more variable than responses to SSRIs (ie, the reference group);

Table. Results of Secondary Analyses Examining Baseline Depression Severity, and Randomized Clinical Trial Publication Year

Factor	b_{ln} $\hat{o}$ (95% CI)[a]	*P* value
Baseline depression		
Baseline depression × AD	0.99 (0.96-1.02)	.58
Baseline depression mean	1.08 (1.04-1.12)	<.001
AD (compared with placebo)	1.00 (0.97-1.03)	.83
End point mean	1.04 (1.00-1.07)	.03
Publication year		
Publication year × AD	0.999 (0.996-1.002)	.42
Publication year	0.998 (0.994-1.002)	.36
AD (compared with placebo)	1.040 (1.000-1.079)	.048
End point mean	1.081 (1.040-1.116)	<.001

Abbreviation: AD, antidepressant.

[a] b_{ln} $\hat{o}$ reflects the association of each predictor with *ln* $\hat{o}$.

however, responses to NAs were 11% more variable than responses to SSRIs (b_{ln} $\hat{o}$, 1.11; 95% CI, 1.01-1.21; P = .02) (**Figure 3**).

RCT Publication Year

The 71 RCTs with a publication year available, corresponding to 158 observations (Figure 1), were published between 1979 and 2018. We found no evidence that the association between treatment group and variability association with the year that an RCT was published (b_{ln} $\hat{o}$ = 0.999; 95% CI, 0.996-1.002; P = .42) (see the Table for model predictor results).

Sensitivity Analyses

We examined whether results from our primary and secondary analyses were affected by depression scales used to measure end point means and SDs across RCTs. Results were not qualitatively different when controlling for scale (full results are provided in eTables 2-7 in the Supplement). One difference emerged when controlling for treatment duration: the difference in variability between SSRIs and NAs was qualitatively similar to our secondary analysis, but it was no longer statistically significant (b_{ln} $\hat{o}$, 1.108; 95% CI, 0.997-1.227; P = .05) (eTables 8-12 in the Supplement). Results from our primary and secondary analyses were not qualitatively different in RSMMs using expanded priors (eTables 13-17 in the Supplement).

Discussion

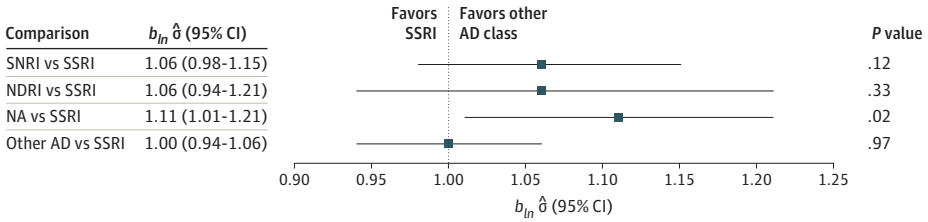
We examined whether there is evidence of systematic variability in the overall symptomatic response to antidepressants by comparing variability in outcomes in RCTs between participants with MDD assigned to receive antidepressants or placebo. However, contrary to our hypothesis, we found no evidence that variability in observed response (as measured by total depression scores) among participants receiving antidepressants was greater than among those receiving a placebo. Our results do not identify variability based on total depression scores. In other words, our findings do not provide empirical support for efforts to personalize antidepressant treatment based on total depression scores. Nevertheless, these findings do not rule out the possibility of treatment effect heterogeneity; they are consistent with a range of treatment effect SDs that are greater than zero (up to 16.10), depending on negative associations between the treatment and placebo effects (eTable 1 in the Supplement). These findings also do not address the potential for personalization based on individual symptoms or biomarkers, where it may be possible to select specific antidepressants according to some specific clinical or biological characteristics of individual patients.[37]

Our findings are consistent with other work suggesting that random or nonspecific factors may account for the variability in observed response to antidepressants. In these studies, variability ratios (VRs) were calculated to quantify differences in SDs of pre/post differences in depression between antidepressant and placebo groups in an earlier open data set.[21] Consistent with our results, meta-analyses of VRs[13,19,20] found no evidence of larger variability in antidepressant groups as compared with placebo. Our findings are also consistent with previous work in another field of psychiatry,[6] which showed that overall responses to antipsychotics are not more variable than responses to placebo in patients with schizophrenia.

Findings from our secondary analyses suggest that traditional MDD subtypes based on symptom severity are not associated with variability in observed response. This result is consistent with other studies showing that the efficacy of antidepressants is comparable across the continuum of symptom severity of MDD.[38-40] Coupled with our primary finding, these results do not support the assumption that there are moderators of observed responses to antidepressants associated with overall severity of depression. At the same time, if MDD is a heterogeneous condition with diverse symptom profiles or subtypes, total depression scores may not adequately capture its typology or even its severity. Some have argued that examining responses to antidepressants based on total scores conceals their effects,[41,42] which may also apply to the variability in these effects. Although our results do not provide evidence in support of personalizing MDD treatments based on total depression scores, they do not rule out the possibility of selecting antidepressant treatments based on scores corresponding to individual symptoms or biomarkers. This is consistent with the suggestion that a focus on total depression scores may thwart progress in personalizing treatments for MDD, since there may not be individual differences or biomarkers that capture the shared variance of all its symptoms.[42] The next step toward optimizing treatments should involve examining the evidence for variability in responses to antidepressants based on individual symptoms (eg, suicidality[43]), symptom profiles,[44] sets of biomarkers, or a combination of them.

In our secondary analysis, we did not find any differences in the variability of responses to antidepressant classes and placebo. However, a difference emerged among antidepressant classes, with responses to SSRIs being less variable than responses to NAs. In our sensitivity analysis, this association may be owing to longer treatment duration with NAs. Nevertheless, responses to NAs may be more variable than responses to SSRIs because antidepressants affecting primarily norepinephrine (eg, amitriptyline) may have a greater effect

Figure 3. Comparisons of Variability Between Selective Serotonin Reuptake Inhibitors (SSRIs) and Antidepressant (AD) Classes



Comparisons of SSRI (the reference group) and AD classes indicate the difference between the two groups in ln $\hat{\sigma}$. These estimates have been exponentiated, with a $b_{ln}\hat{\sigma}$ less than 1 representing less variability than SSRI. NDRI indicates norepinephrine-dopamine reuptake inhibitor; NA, norepinephrinergic agent; SNRI, serotonin and norepinephrine reuptake inhibitor.

on depressive symptoms than antidepressants that affect only synaptic serotonin.[45-48] Alternatively, functional unblinding in RCTs of NAs may account for this finding. Cipriani et al[21] reported that, compared with SSRIs, dropout rates owing to adverse effects were generally higher for antidepressants with noradrenergic effects, introducing the possibility that raters in some RCTs were unblinded to treatment allocation. Nevertheless, our finding does not provide insight into the association of individual differences in response to NAs compared with placebo, but it suggests that, among the antidepressant classes we examined, NAs may be most amenable to personalization.

Limitations

The results of our study should be considered in the context of some methodologic challenges and limitations. The analysis of variability in response to psychotropic medications has attracted a renewed interest with the availability of novel methods to summarize this variability.[18] To our knowledge, the first study to use one of these methods in relation to psychotropic medications was published in 2019[6]; since then, more than 5 published studies[13,14,17,19,20] (including this analysis) and other documents[15,16] have described the merits and shortcomings of these approaches. When comparing variability between antidepressant and placebo groups, coefficient of variation ratios (CVRs) have been used to account for associations between end point scores and SDs. However, CVRs can yield biased results depending on the nature of these associations, specifically by inflating variability in the group with a lower end point score.[13,17,18] Previous findings of more variability in response to placebo as compared with antidepressant derived with the CVR might therefore be associated with a focus in these studies on pre/post difference scores,[19,20] which tend to be lower in placebo groups. Owing to the nature of the mean-SD association in our data, the use of VRs and CVRs could produce biased results (as described in the eMethods and depicted in eFigures 1 and 2 in the Supplement). Given that similar investigations of variability in response to other psychiatric treatments and interventions are emerging,[49] it is crucial to consider when the VR and CVR are not appropriate and when mixed models that control for the effect of the mean on variability, such as the RSMM, are optimal.[13] Also, our findings do not rule out heterogeneity in treatment effects (ie, the difference in outcomes if the same individuals received antidepressant and placebo). According to a 2020 analysis,[7] a finding of increased variability in response to treatment suggests treatment effect heterogeneity. Adapting this analysis[7] in the context of our study, we show that our finding of no increased variability in response to antidepressant is consistent with a range of potential treatment effect heterogeneities, up to twice the size of the SD in those treated with placebo (described in the eResults in the Supplement).

The validity of our results rests on the quality of RCTs that were included in our data set, which in some cases was low.[21] Because participants in RCTs are selected to be relatively homogenous, our results may also not generalize to patients with MDD seen in clinical practice; it is possible there may be more variability in overall response to antidepressants in patients with greater variability in clinical characteristics. Furthermore, our analysis was limited to RCTs with available data at end point, as well as the 15 drugs from the eligible RCTs. It is possible that some antidepressants that we did not include would produce different results because of their specific mechanisms of action. We grouped antidepressants into classes; therefore, it is likely that our results apply to other antidepressants belonging to the same class. However, few comparisons were available for SNRIs and NDRIs, and our analyses involving these antidepressant classes may have been underpowered.

Given that we relied on published data, we were limited in the moderators of response variability that we could address. Without access to patient-level or item-level data, we could not examine how various subgroups of patients or symptom profiles affected response variability. This remains an important direction for future work. As discussed in previous work,[6,15,16] our analysis does not rule out the possibility of treatment by subgroup interactions or address the potential to personalize treatments based on subsets of symptoms. Finally, our findings also do not address variability in response to nonpharmacological treatments, such as psychotherapies, convulsive therapies, and neuromodulatory interventions, which might possibly be personalized based on their effects.

Conclusions

In conclusion, our results do not support the widely held assumption that individual differences underlie the variability

in the association between total depression scores and antidepressant treatment. Future efforts toward personalizing treatments for MDD should focus on whether individual symptoms or biomarkers are associated with variability in response to antidepressants or whether there is evidence of variability in response to other treatment types.

ARTICLE INFORMATION

Accepted for Publication: November 28, 2020.

Published Online: February 17, 2021. doi:10.1001/jamapsychiatry.2020.4564

Author Affiliations: Centre for Addiction and Mental Health, Toronto, Ontario, Canada (Maslej, Sanches, Mulsant); Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada (Maslej, Mulsant); Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine, School of Public Health, Yoshida-Konoe, Sakyo, Kyoto, Japan (Furukawa); Department of Clinical Epidemiology, Kyoto University Graduate School of Medicine, Kyoto University School of Public Health, Yoshida-Konoe, Sakyo, Kyoto, Japan (Furukawa); Department of Psychiatry, University of Oxford, Oxford, England (Cipriani, Tomlinson); Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, England (Cipriani, Tomlinson); Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, Canada (Andrews); Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin Berlin, Berlin, Germany (Volkmann); Institute of Psychiatry, Psychology and Neuroscience, Department of Psychosis Studies, King's College of London, London, England (McCutcheon, Howes, Guo); Department of Psychiatry, Renmin Hospital of Wuhan University, Wuhan, China (Guo).

Author Contributions: Dr Maslej had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Maslej, Cipriani, Howes, Mulsant.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Maslej, Andrews, Tomlinson, McCutcheon, Mulsant.
Critical revision of the manuscript for important intellectual content: Furukawa, Cipriani, Andrews, Sanches, Tomlinson, Volkmann, McCutcheon, Howes, Guo, Mulsant.
Statistical analysis: Maslej, Sanches, Volkmann, McCutcheon, Guo.
Obtained funding: Maslej, Cipriani.
Administrative, technical, or material support: Tomlinson.
Supervision: Cipriani, Sanches, Howes, Mulsant.

Conflict of Interest Disclosures: Dr Furukawa reported receiving honoraria for lectures from MSD and consulting fees from Mitsubishi-Tanabe and Shionogi; receiving a grant from Mitsubishi-Tanabe outside the submitted work; and having patent 2018-177688 pending and a copyright concerning Kokoro-app issued. Dr Cipriani has received research and consultancy fees from INCiPiT (Italian Network for Paediatric Trials), CARIPLO Foundation, and Angelini Pharma. Dr Tomlinson has received research and consultancy fees from INCiPiT (Italian Network for Paediatric Trials) and Angelini Pharma. Dr Howes has received investigator-initiated research funding from and/or participated in advisory/speaker meetings organised by Angelini, AstraZeneca, Autifony, Biogen, Boehringer Ingelheim, Eli Lilly, Heptares, Invicro, Janssen, Lundbeck, Lyden-Delta, Mylan, Neurocrine, Otsuka, Sunovion, Rand, Recordati, and Roche. Neither Dr Howes or his family have been employed by or have holdings/a financial stake in any pharmaceutical company. Dr Mulsant reported currently receiving research support from Brain Canada, the Canadian Institutes of Health Research, the CAMH Foundation, the Patient-Centered Outcomes Research Institute, the National Institutes of Health (NIH), Capital Solution Design LLC (software used in a study founded by the CAMH Foundation), and HAPPYneuron (software used in a study founded by Brain Canada); owning stock in General Electric (less than $5000); and having received research support from Eli Lilly (medications for an NIH-funded clinical trial) and Pfizer (medications for a NIH-funded clinical trial). Dr Mulsant holds and receives support from the Labatt Family Chair in Biology of Depression in Late-Life Adults at the University of Toronto. No other disclosures were reported.

Funding/Support: Dr Maslej is funded by the CIHR Health Systems Impact Postdoctoral Fellowship. Dr Cipriani is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017-08-ST2-006), by the NIHR Applied Research Collaboration Oxford and Thames Valley at Oxford Health NHS Foundation Trust, and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005). Dr McCutcheon is funded by the NIHR. Dr Howes is funded by Medical Research Council-UK (MC U120097115), Maudsley Charity (666), Brain and Behavior Research Foundation, and Wellcome Trust (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. Dr Guo is funded by the National Key R&D Program of China (2018YFC1314600).

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The views expressed are those of the authors and not necessarily those of the UK National Health Service, National Institute for Health Research, or UK Department of Health and Social Care.

Additional Contributions: We thank Cipriani et al[21] for making data from their network meta-analysis available online and Nakagawa et al[18] for making their code available online. We also thank an anonymous, uncompensated reviewer for suggesting the secondary analysis involving RCT publication year. We thank Benjamin M. Bolker, PhD, Departments of Biology and Mathematics & Statistics, McMaster University, Hamilton, Ontario, Canada, for reviewing our primary analysis and providing his edits and suggestions regarding our code.

REFERENCES

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association; 2013.

2. National Institute for Health and Care Excellence. Depression in adults: recognition and management: clinical guideline. Published October 28, 2009. Accessed January 15, 2021. https://www.nice.org.uk/guidance/cg90/resources/depression-in-adults-recognition-and-management-pdf-975742636741

3. Thomas L, Kessler D, Campbell J, et al. Prevalence of treatment-resistant depression in primary care: cross-sectional data. *Br J Gen Pract*. 2013;63(617):e852-e858. doi:10.3399/bjgp13X675430

4. Noma H, Furukawa TA, Maruo K, et al. Exploratory analyses of effect modifiers in the antidepressant treatment of major depression: individual-participant data meta-analysis of 2803 participants in seven placebo-controlled randomized trials. *J Affect Disord*. 2019;250:419-424. doi:10.1016/j.jad.2019.03.031

5. Uher R. Genes, environment, and individual differences in responding to treatment for depression. *Harv Rev Psychiatry*. 2011;19(3):109-124. doi:10.3109/10673229.2011.586551

6. Winkelbeiner S, Leucht S, Kane JM, Homan P. Evaluation of differences in individual treatment response in schizophrenia spectrum disorders: a meta-analysis [published online June 3, 2019]. *JAMA Psychiatry*. 2019;76(10):1063-1073. doi:10.1001/jamapsychiatry.2019.1530

7. Volkmann A. On the relationship between treatment effect heterogeneity and the variability ratio effect size statistic. *arXiv*. Accessed January 15, 2021. https://arxiv.org/pdf/2006.11848.pdf

8. Senn S. Mastering variation: variance components and personalised medicine. *Stat Med*. 2016;35(7):966-977. doi:10.1002/sim.6739

9. Cuijpers P, Christensen H. Are personalised treatments of adult depression finally within reach? *Epidemiol Psychiatr Sci*. 2017;26(1):40-42. doi:10.1017/S204579601600007X

10. Kessler RC. The potential of predictive analytics to provide clinical decision support in depression treatment planning. *Curr Opin Psychiatry*. 2018;31(1):32-39. doi:10.1097/YCO.0000000000000377

11. Ozomaro U, Wahlestedt C, Nemeroff CB. Personalized medicine in psychiatry: problems and promises. *BMC Med*. 2013;11:132. doi:10.1186/1741-7015-11-132

12. Simon GE, Perlis RH. Personalized medicine for depression: can we match patients with treatments? *Am J Psychiatry*. 2010;167(12):1445-1455. doi:10.1176/appi.ajp.2010.09111680

13. Volkmann C, Volkmann A, Müller CA. On the treatment effect heterogeneity of antidepressants in major depression: a Bayesian meta-analysis and simulation study. *PLoS One* doi:10.1371/journal.pone.0241497

14. Mills HL, Higgins JP, Morris RW, et al. Detecting heterogeneity of intervention effects using analysis and meta-analysis of differences in variance between arms of a trial. *medRxiv*. Accessed May 12, 2020. doi:10.1101/2020.03.07.20032516

15. Winkelbeiner S, Homan P. Is variance ratio a valid indicator of heterogeneous treatment effect? reply. *JAMA Psychiatry*. 2020;77(2):217-218. doi:10.1001/jamapsychiatry.2019.3382

16. Bae S. Is variance ratio a valid indicator of heterogeneous treatment effect? *JAMA Psychiatry*. 2020;77(2):216-217. doi:10.1001/jamapsychiatry.2019.3372

17. Guo X, McCutcheon RA, Pillinger T, et al. The magnitude and heterogeneity of antidepressant response in depression: A meta-analysis of over 45,000 patients. *J Affect Disord*. 2020;276:991-1000. doi:10.1016/j.jad.2020.07.102

18. Nakagawa S, Poulin R, Mengersen K, et al Meta-analysis of variation: ecological and evolutionary applications and beyond. *Methods Ecol Evol*. 2015;6(2):143-152. doi:10.1111/2041-210X.12309

19. Plöderl M, Hengartner MP. What are the chances for personalised treatment with antidepressants? detection of patient-by-treatment interaction with a variance ratio meta-analysis. *BMJ Open*. 2019;9(12):e034816. doi:10.1136/bmjopen-2019-034816

20. Munkholm K, Winkelbeiner S, Homan P. Individual response to antidepressants for depression in adults–a simulation study and meta-analysis. *PloS One*. 2020;15(8):e0237950.

21. Cipriani A, Furukawa TA, Salanti G, et al. 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*. 2018;391(10128):1357-1366. doi:10.1016/S0140-6736(17)32802-7

22. Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA*. 2010;303(1):47-53. doi:10.1001/jama.2009.1943

23. Stone M, Kalaria S, Richardville K, et al Components and trends in treatment effects in randomized placebo-controlled trials in major depressive disorder from 1979-2016. Presented at: American Society of Clinical Psychopharmacology Annual Conference; May 30, 2018; Miami, FL.

24. Furukawa TA, Salanti G, Atkinson LZ, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *BMJ Open*. 2016;6(7):e010919. doi:10.1136/bmjopen-2015-010919

25. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23(1):56-62. doi:10.1136/jnnp.23.1.56

26. Bech P. *Clinical Psychometrics*. Oxford, England: Wiley Blackwell; 2012. doi:10.1002/9781118511800

27. Williams JBW, Link MJ, Rosenthal NE, Terman M. *Structured Interview Guide for the Hamilton Depression Rating Scale, Seasonal Affective Disorders Version (SIGH-SAD)*. New York, New York: State Psychiatric Institute, Biometrics Research; 1988.

28. Roberts JN, Lei HH, Krishen A, et al Identifying sub-scales of the 31-item Hamilton Depression Scale. Poster presented at: New Clinical Drug Evaluation Unit (NCDEU) meeting; May 28-31, 2001; Phoenix, AZ.

29. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382-389. doi:10.1192/bjp.134.4.382

30. Rush AJ, Giles DE, Schlesser MA, Fulton CL, Weissenburger J, Burns C. The Inventory for Depressive Symptomatology (IDS): preliminary findings. *Psychiatry Res*. 1986;18(1):65-87. doi:10.1016/0165-1781(86)90060-0

31. Raudenbush SW, Bryk AS. Examining correlates of diversity. *J Educ Stat*. 1987; 12(3):241-69. doi:10.3102/10769986012003241

32. Ghaemi SN. A new nomenclature for psychotropic drugs. *J Clin Psychopharmacol*. 2015;35(4):428-433. doi:10.1097/JCP.0000000000000341

33. Owens MJ, Morgan WN, Plott SJ, Nemeroff CB. Neurotransmitter receptor and transporter binding profile of antidepressants and their metabolites. *J Pharmacol Exp Ther*. 1997;283(3):1305-1322.

34. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Published 2019. Accessed January 12, 2020. https://www.R-project.org/

35. Hadfield J. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Softw*. 2010;33(2):1-22. doi:10.18637/jss.v033.i02

36. Open Science Foundation. Individual differences in response to antidepressants: R Code. Accessed November 13, 2020. https://mfr.ca-1.osf.io/render?url=https://osf.io/m6u89/?direct%26mode=render%26action=download%26mode=render

37. Chekroud AM, Gueorguieva R, Krumholz HM, Trivedi MH, Krystal JH, McCarthy G. Reevaluating the efficacy and predictability of antidepressant treatments: a symptom clustering approach. *JAMA Psychiatry*. 2017;74(4):370-378. doi:10.1001/jamapsychiatry.2017.0025

38. Furukawa TA, Maruo K, Noma H, et al. Initial severity of major depression and efficacy of new generation antidepressants: individual participant data meta-analysis. *Acta Psychiatr Scand*. 2018;137(6):450-458. doi:10.1111/acps.12886

39. Rabinowitz J, Werbeloff N, Mandel FS, Menard F, Marangell L, Kapur S. Initial depression severity and response to antidepressants v. placebo: patient-level data analysis from 34 randomised controlled trials. *Br J Psychiatry*. 2016;209(5):427-428. doi:10.1192/bjp.bp.115.173906

40. Gibbons RD, Hur K, Brown CH, Davis JM, Mann JJ. Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. *Arch Gen Psychiatry*. 2012;69(6):572-579. doi:10.1001/archgenpsychiatry.2011.2044

41. Fried EI, Nesse RM. Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *J Affect Disord*. 2015;172:96-102. doi:10.1016/j.jad.2014.10.010

42. Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Med*. 2015;13(1):72. doi:10.1186/s12916-015-0325-4

43. Niculescu AB, Le-Niculescu H, Levey DF, et al. Precision medicine for suicidality: from universality to subtypes and personalization. *Mol Psychiatry*. 2017;22(9):1250-1273. doi:10.1038/mp.2017.128

44. Green E, Goldstein-Piekarski AN, Schatzberg AF, Rush AJ, Ma J, Williams L. Personalizing antidepressant choice by sex, body mass index, and symptom profile: An iSPOT-D report. *Personalized Medicine in Psychiatry*. 2017;1:65-73. doi:10.1016/j.pmip.2016.12.001

45. Picardi A, Viroli C, Tarsitani L, et al. Heterogeneity and symptom structure of schizophrenia. *Psychiatry Res*. 2012;198(3):386-394. doi:10.1016/j.psychres.2011.12.051

46. Tandon R, Gaebel W, Barch DM, et al. Definition and description of schizophrenia in the DSM-5. *Schizophr Res*. 2013;150(1):3-10. doi:10.1016/j.schres.2013.05.028

47. Durisko Z, Mulsant BH, Andrews PW. An adaptationist perspective on the etiology of depression. *J Affect Disord*. 2015;172:315-323. doi:10.1016/j.jad.2014.09.032

48. Thase ME, Entsuah AR, Rudolph RL. Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *Br J Psychiatry*. 2001;178(3):234-241. doi:10.1192/bjp.178.3.234

49. Radua J, Davies C, Fusar-Poli P. Evaluation of variability in individual response to treatments in the clinical high-risk state for psychosis: a meta-analysis. *Schizophr Res*.S0920-9964(20)30262-0. doi:10.1016/j.schres.2020.05.010