

## Original Article

# Moderators of antidepressant augmentation versus switch in the OPTIMUM randomised controlled trial

Helena K. Kim, Jordan F. Karp\*, Helen Lavretsky\*, Daniel M. Blumberger\*\*, Patrick J. Brown\*\*, Alastair J. Flint\*\*, Emily Lenard\*\*, J. Philip Miller\*\*, Charles F. Reynolds\*\*, Steven P. Roose\*\*, Eric J. Lenze\*\*\* and Benoit H. Mulsant\*\*\*

## Background

Older adults with treatment-resistant depression (TRD) benefit more from treatment augmentation than switching. It is useful to identify moderators that influence these treatment strategies for personalised medicine.

## Aims

Our objective was to test whether age, executive dysfunction, comorbid medical burden, comorbid anxiety or the number of previous adequate antidepressant trials could moderate the superiority of augmentation over switching. A significant moderator would influence the differential effect of augmentation versus switching on treatment outcomes.

## Method

We performed a preplanned moderation analysis of data from the Optimizing Outcomes of Treatment-Resistant Depression in Older Adults (OPTIMUM) randomised controlled trial ( $N = 742$ ). Participants were 60 years old or older with TRD. Participants were either (a) randomised to antidepressant augmentation with aripiprazole (2.5–15 mg), bupropion (150–450 mg) or lithium (target serum drug level 0.6 mmol/L) or (b) switched to bupropion (150–450 mg) or nortriptyline (target serum drug level 80–120 ng/mL). Treatment duration was 10 weeks. The two main outcomes of this analysis were (a) symptom improvement, defined as change in Montgomery–Åsberg Depression Rating Scale (MADRS) scores from baseline to week 10 and (b) remission, defined as MADRS score of 10 or less at week 10.

## Results

Of the 742 participants, 480 were randomised to augmentation and 262 to switching. The number of adequate previous antidepressant trials was a significant moderator of depression symptom improvement ( $b = -1.6$ ,  $t = -2.1$ ,  $P = 0.033$ , 95% CI  $[-3.0, -0.1]$ , where  $b$  is the coefficient of the relationship (i.e. effect size), and  $t$  is the  $t$ -statistic for that coefficient associated with the  $P$ -value). The effect was similar across all augmentation strategies. No other putative moderators were significant.

## Conclusions

Augmenting was superior to switching antidepressants only in older patients with fewer than three previous antidepressant trials. This suggests that other intervention strategies should be considered following three or more trials.

## Keywords

Treatment-resistant depression; late-life depression; antidepressant; augmentation; moderator.

## Copyright and usage

© The Author(s), 2025. Published by Cambridge University Press on behalf of Royal College of Psychiatrists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Treatment-resistant depression (TRD) is typically defined as a major depressive disorder that does not remit after at least two antidepressant trials of adequate dosage and duration.<sup>1</sup> Common pharmacotherapy strategies for patients with TRD include augmentation with another medication (which could be a second antidepressant or a psychotropic medication from another class – for example, an atypical antipsychotic or a mood stabiliser) or switching to a different antidepressant. The Optimizing Outcomes of Treatment-Resistant Depression in Older Adults (OPTIMUM) clinical trial compared these two strategies in older adults with TRD.<sup>2,3</sup> We recently reported that augmentation with aripiprazole is superior to switching to bupropion,<sup>3</sup> consistent with a growing body of literature favouring augmentation over switching strategies in the treatment of TRD.<sup>4,5</sup> In this context, the TRD field has converged on comparing these two strategies rather than looking at specific medications.<sup>4</sup>

Identifying moderators that influence the effectiveness of these treatment strategies advances personalised medicine. For example, a moderator that can be assessed by a clinician, such as age, helps the clinician choose the optimal treatment strategy based on

individual characteristics in day-to-day practice. Moderators can also expand our understanding of the biological mechanisms involved in a disorder or its treatment. To date, five factors have consistently been shown to negatively influence treatment outcomes with antidepressant treatment in late life: age;<sup>6</sup> executive dysfunction;<sup>7</sup> comorbid medical burden;<sup>8,9</sup> comorbid anxiety;<sup>6,7</sup> and degree of treatment resistance.<sup>10,11</sup> Therefore, the second aim of the OPTIMUM study was to assess whether these five factors moderate the effect of augmentation versus switching. A preplanned moderation analysis was performed to test the hypothesis that these five factors will moderate symptom improvement and depression remission in older adults with TRD. Baseline depression severity, which can bias the interpretation of factors associated with treatment outcomes,<sup>12</sup> was added as a covariate in the moderation analysis, as done in previous analyses of clinical trials.<sup>13–15</sup>

## Method

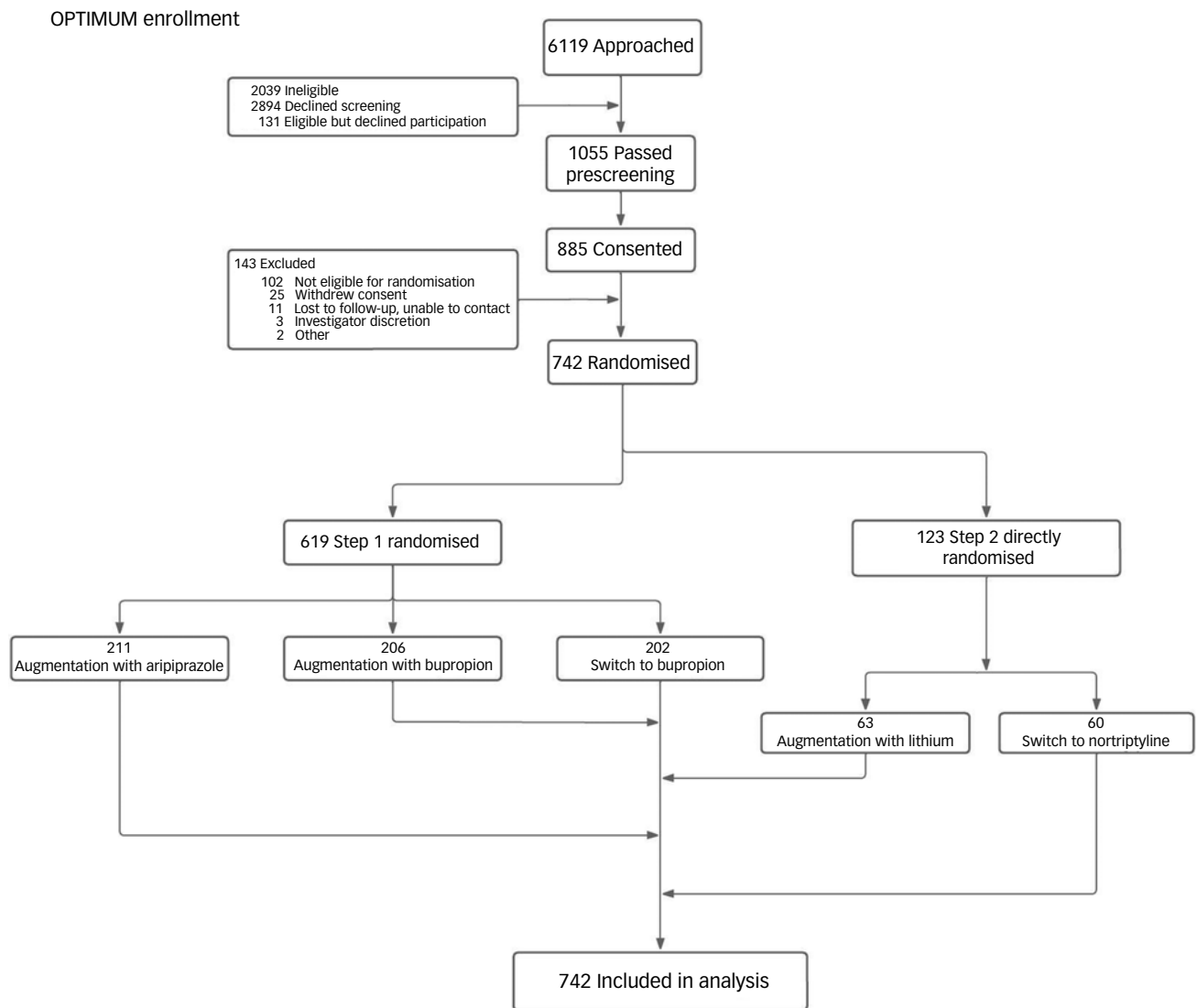
### Study overview

OPTIMUM was a pragmatic, open-label, randomised clinical trial (NCT02960763) that received approval by the institutional review boards of the five sites where the trial was conducted (Columbia

\* These authors contributed equally and are listed in alphabetical order.

\*\* These authors contributed equally and are listed in alphabetical order.

\*\*\* These authors contributed equally and are listed in alphabetical order.



**Fig. 1** Enrollment, randomisation and inclusion in analysis from steps 1 and 2. OPTIMUM, Optimizing Outcomes of Treatment-Resistant Depression in Older Adults.

University; University of California, Los Angeles; University of Pittsburgh; University of Toronto; and Washington University in St. Louis). Its design,<sup>2</sup> sample size calculation, recruitment, follow-up<sup>3</sup> and primary report<sup>3</sup> have been published previously.

### Ethics statement

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation, and with the Helsinki Declaration of 1975 as revised in 2013. All procedures involving human subjects/patients were approved by the institutional review boards (IRBs) at each of the five trial sites (IRB ID nos: Centre for Addiction and Mental Health, 086/2016; New York State Psychiatric Institute/Columbia University, 7409; University of Pittsburgh, PRO16100179; University of California, Los Angeles, 16-001829; Washington University in St. Louis, 201609085).

### Randomisation

Details of randomisation have been published previously.<sup>3</sup> Briefly, in step 1 (duration, 10 weeks) of the trial, participants were randomly assigned in a 1:1:1 ratio to augmenting their current

antidepressant with aripiprazole (2.5–15 mg daily) or bupropion (150–450 mg daily) or switching to bupropion (150–450 mg). If the participant had already tried one of the step 1 medications without clinical benefit or tolerability, they entered step 2 directly without going through step 1 (duration, 10 weeks) and were randomised in a 1:1 ratio to either (a) augmentation of their current antidepressant with lithium (targeting a serum drug level of 0.6 mmol/L) or (b) a switch to nortriptyline (targeting a serum drug level of 80–120 ng/mL). In this preplanned moderation analysis, we used data from participants in step 1 and from those who entered directly into step 2 (Fig. 1). We did not include participants who were entered into step 2 after not attaining remission in step 1. This is because those who entered step 2 after not attaining remission in step 1 would have had 10 additional weeks of treatment in a clinical trial setting compared with those who were in step 1, or compared with those who had direct entry into step 2. This approach permitted us to analyse baseline and week 10 data (i.e. 10 weeks after they had started their randomised pharmacotherapy).

### Participant characteristics

Participants were 60 years of age or older with TRD, defined as not remitting after two or more antidepressant trials of adequate dosage

and duration at any time during the current episode. TRD was ascertained using the Antidepressant Treatment History Form (ATHF; Buchalter et al<sup>11</sup>). Race was self-reported by the participants. All participants gave written informed consent after being fully informed about the study details.

## Statistical analysis

The published primary report of the OPTIMUM results showed the superiority of aripiprazole augmentation compared with switching to bupropion for change in depression severity and remission.<sup>3</sup> Because the sample size was adjusted from 1500 to 708 mid-study, a post hoc calculation showed that a sample size of 561 would provide sufficient power ( $0.8, \alpha = 0.05$ ) to detect a small effect size ( $0.14$ ) in a regression model consisting of 3 variables (treatment arm, moderator and covariate). There is robust evidence supporting the superiority of augmentation over switch strategies, with the TRD field converging on comparing these two strategies rather than looking at the specific medications.<sup>4,5</sup> Therefore, our statistical analysis plan specified that this preplanned analysis would focus on comparing the overall strategy of augmentation versus switch on the same two depression clinical outcomes as the primary report (i.e. change in depression severity and remission). The five *a priori* hypothesised moderators were age, executive dysfunction, comorbid medical burden, comorbid anxiety and degree of treatment resistance. Executive dysfunction was measured using the US National Institutes of Health (NIH) Toolbox fluid cognition composite, with a score of 85 or lower indicating impairment.<sup>16</sup> Comorbid medical burden was measured using the Cumulative Illness Rating Scale – Geriatric (CIRS-G).<sup>17</sup> Comorbid anxiety was measured using the Patient Reported Outcomes Measurement Information System (PROMIS) anxiety scale, with a score of 60 or higher (one standard deviation above average anxiety) being considered clinically anxious.<sup>18</sup> The degree of treatment resistance was defined as the number of previous adequate antidepressant trials during the past 2 years, or during the current episode if the episode had lasted less than 2 years. The number of previous adequate antidepressant trials during the past 2 years (or current episode if the episode had lasted less than 2 years) was obtained from ATHF.<sup>19</sup> Only antidepressants prescribed during either the duration of the current major depressive episode or the past 2 years, whichever was shorter, were considered when rating ATHF, because of concerns over the accuracy of recalling specific antidepressants taken years before (and even more so, their dosages and duration). Executive dysfunction and comorbid anxiety were dichotomised because the scales used to measure these hypothesised moderators have thresholds for determining clinically significant impairment, allowing for clinical application of the findings. Based on the results of the primary report,<sup>3</sup> we anticipated that approximately 15% of participants included in this analysis would not have week 10 MADRS scores. The two outcomes were change in MADRS scores (i.e. depression symptom improvement) and remission. We used the same definition of remission as in the published primary report,<sup>3</sup> which was defined as MADRS<sup>20</sup> score  $\leq 10$  at week 10 (of step 1 or 2). MADRS  $< 10$  was reported to be concordant with remission as defined by the Clinical Global Impression Scale for Severity.<sup>21</sup> We performed Little's missing completely at random (MCAR) test:<sup>22</sup> as expected with an effectiveness clinical trial, the outcome data were missing not at random (MNAR;  $\chi^2 = 1115.9$ , d.f. = 75,  $P = 0.002$ ). Thus, we did not impute MNAR missing outcome data.

IBM SPSS 27 for Windows was used for all statistical analyses, with averages expressed as mean  $\pm$  standard deviation. Analysis of covariance (ANCOVA) was used for comparing change in MADRS scores between the two treatment strategies. Log binomial

regression was used for comparing remission rates between the two treatment strategies, because it carries a lower risk of inflating effect sizes and produces more readily interpretable outcomes than logistic regression.<sup>23</sup> The results are expressed in prevalence ratios and confidence intervals. Residual plots were used to check for homoscedasticity (i.e. assumption of equal variance). The PROCESS Macro extension<sup>24</sup> was used to perform moderation analysis through previously reported methods.<sup>25</sup> Moderation analysis was performed with listwise deletion to handle missing data as per the design of the programme. Briefly, moderation analysis using PROCESS performs a regression analysis with the independent variable (i.e. treatment strategies), the potential moderator (e.g. age) and the interaction term between the independent variable and the potential moderator predicting the dependent variable (e.g. change in MADRS scores). Bootstrapping is performed in PROCESS to mitigate non-normal distribution of data. A moderation effect is found if the interaction term has a significant effect on the dependent variable. Conditional effects of the independent variable on the dependent variable are also given at different levels of the moderator, to illustrate the moderating effect. Continuous variables that define products were mean centred. Simple slope analysis was performed for moderators identified as being significant, with each potential moderator being analysed separately. Baseline MADRS score was included as a covariate in all analyses. A Spearman's rank correlation analysis was performed among the five potential moderators and baseline MADRS scores. Age ( $\rho = -0.10$ ,  $n = 717$ ,  $P = 0.008$ ), comorbid anxiety ( $\rho = 0.3$ ,  $n = 623$ ,  $P < 0.001$ ) and number of previous adequate antidepressant trials ( $\rho = 0.08$ ,  $n = 717$ ,  $P = 0.042$ ) were all significantly correlated with baseline MADRS scores, supporting the addition of baseline MADRS as a covariate. A further Spearman's rank correlation analysis was performed among the five potential moderators to identify potential confounders. The only two significant correlations found were between age and CIRS-G score ( $\rho = 0.13$ ,  $P < 0.001$ ) and between age and executive dysfunction ( $\rho = 0.31$ ,  $P < 0.001$ ). Subsequently, moderation analysis for age included CIRS-G score and executive dysfunction as covariates; moderation analysis for both CIRS-G and executive dysfunction included age as a covariate. Corrections for multiple comparisons were not performed, because this preplanned analysis assessed a small set of prespecified potential moderators.<sup>26</sup> We used the previously published minimal clinically important difference (MCID) in MADRS scores of 1.6–1.9<sup>27</sup> to interpret the findings.

## Results

### Participant characteristics

Most of the participants were from step 1 and, by design, around two-thirds received augmentation while the other third were switched to a different agent (Fig. 1). Participant characteristics are shown in Table 1. Of 742 participants, 480 (64.7%) received augmentation (of whom 415 [86.5%] had week 10 MADRS scores) and 262 (35.3%) were switched to a different treatment (of whom 215 [82.1%] had week 10 MADRS scores; total discontinuation rate, 15.1%). Because some data were missing for three moderator variables, the following number (%) of participants had complete data for CIRS-G score: 476 (99.2%) in augmentation, 261 (99.6%) in switch; PROMIS anxiety score: 408 (85.0%) in augmentation, 227 (86.6%) in switch; and fluid cognition score, 295 (61.5%) in augmentation, 168 (69.4%) in switch.

### Moderation analyses

We combined steps 1 and 2 for our analysis.<sup>3</sup> Consistent with the original report, our sample showed a larger mean (s.d.) change in

**Table 1** Participant characteristics

Parameter	Augmentation	Switch	Total <sup>a</sup> (mean ± s.d. or N (%))
Sample size (N)	480	262	742
Steps	Step 1, 417 (86.9%) Step 2, 63 (13.1%)	Step 1, 202 (77.1%) Step 2, 60 (22.9%)	Step 1, 619 (83.4%) Step 2, 123 (16.6%)
Treatment groups			Augmentation with aripiprazole, 211 (28.4%) Augmentation with bupropion, 206 (27.8%) Switch to bupropion, 202 (27.2%) Augmentation with lithium, 63 (8.5%) Switch to nortriptyline, 60 (8.1%)
Treatment strategy			Augmentation, 480 (64.7%) Switch, 262 (35.3%)
Age (years)	68.9 ± 6.6	69.3 ± 7.3	69.1 ± 6.8
Self-reported gender	Women, 333 (69.4%) Men, 147 (30.6%)	Women, 167 (63.7%) Men, 95 (36.3%)	Women, 500 (67.4%) Men, 242 (32.6%)
Race	Caucasian, 408 (85.0%) Other, 72 (15.0%)	Caucasian, 230 (87.8%) Other, 32 (12.2%)	Caucasian, 638 (86.0%) Other, 104 (14.0%)
Years of education	14.6 ± 2.9 ( <i>n</i> = 471)	15.0 ± 2.8 ( <i>n</i> = 257)	14.7 ± 2.9 ( <i>n</i> = 728)
CIRS-G score	8.7 ± 4.7 ( <i>n</i> = 476)	8.6 ± 4.6 ( <i>n</i> = 261)	8.7 ± 4.7 ( <i>n</i> = 737)
Number of previous adequate antidepressant trials (ATHF score)	2.3 ± 0.8 ( <i>n</i> = 480)	2.4 ± 0.9 ( <i>n</i> = 262)	2.4 ± 0.8 ( <i>n</i> = 742)
PROMIS anxiety score	63.6 ± 7.2 ( <i>n</i> = 408)	64.1 ± 7.7 ( <i>n</i> = 227)	63.8 ± 7.4 ( <i>n</i> = 635)
Clinical anxiety	Yes, 303 (74.3%) No, 105 (25.7%)	Yes, 168 (74.0%) No, 59 (26.0%)	Yes, 471 (74.2%) No, 164 (25.8%)
Fluid cognition score	87.0 ± 10.6 ( <i>n</i> = 295)	86.6 ± 11.2 ( <i>n</i> = 168)	86.8 ± 10.8 ( <i>n</i> = 463)
Executive dysfunction	Yes, 140 (47.5%) No, 155 (52.5%)	Yes, 73 (43.5%) No, 95 (56.5%)	Yes, 213 (46.0%) No, 250 (54.0%)
MADRS score at baseline	23.3 ± 7.3 ( <i>n</i> = 464)	23.0 ± 7.0 ( <i>n</i> = 253)	23.2 ± 7.2 ( <i>n</i> = 717)
MADRS score at week 10	16.0 ± 9.2 ( <i>n</i> = 415)	18.1 ± 9.4 ( <i>n</i> = 215)	16.7 ± 9.3 ( <i>n</i> = 630)
Change in MADRS score	7.1 ± 8.0 ( <i>n</i> = 415)	4.7 ± 9.1 ( <i>n</i> = 212)	6.3 ± 8.5 ( <i>n</i> = 627)
Remission at week 10 <sup>b</sup>	Remission, 130 (31.3%) No remission, 285 (68.7%)	Remission, 51 (23.7%) No remission, 164 (76.3%)	Remission, 181 (28.7%) No remission, 449 (71.3%)

ATHF, Antidepressant Treatment History Form; CIRS-G, Cumulative Illness Rating Scale – Geriatric; MADRS, Montgomery–Asberg Depression Rating Scale; PROMIS, Patient-Reported Outcomes Measurement Information System.

a. Total sample size is 742 unless indicated otherwise.

b. Remission defined as having a MADRS score of 10 or less at week 10.

MADRS score with augmentation ( $-7.1 \pm 8.0$ ) than switching ( $-4.7 \pm 9.1$ ). There was also a significant difference in remission rates between the two treatment strategies, where 130 (31.3%) remitted with augmentation versus 51 (23.7%) with switch (prevalence ratio 1.1, 95% CI [1.0, 1.2], *d.f.* = 1,  $P = 0.035$ ).

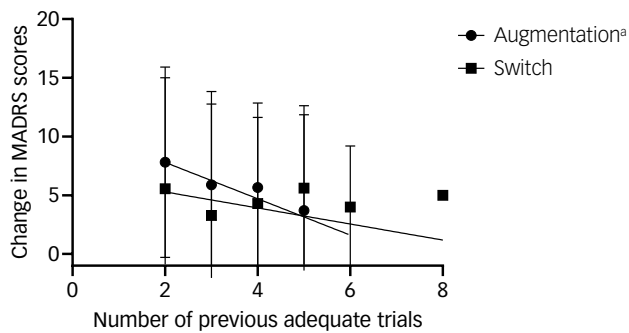
The mean (s.d.) number of adequate previous antidepressant trials was 2.4 (0.8). The number of adequate previous antidepressant trials had a significant negative moderating effect on the change in depression scores ( $b = -1.55$ ,  $t = -2.13$ ,  $P = 0.033$ , 95% CI  $[-2.98, -0.12]$ ), where  $b$  is the coefficient of the relationship (i.e. effect size) and  $t$  is the  $t$ -statistic for that coefficient associated with the  $P$ -value. Examining conditional effects of the focal predictor (i.e. treatment strategies) on the outcome (i.e. change in MADRS score) at different levels of the moderator (i.e. number of previous antidepressant trials) with the Johnson–Neyman technique, using the average as the centre (average number of previous antidepressant trials, 2.4) and comparing those with 1 s.d. below the average versus 1 s.d. above the average,<sup>28</sup> augmentation was superior to switching in participants with 1 s.d. fewer trials compared with the average (effect 3.62,  $P = 0.0001$ ). This superiority of augmentation over switching was diminished in participants with 1 s.d. more previous adequate trials compared with the average (effect 0.88,  $P = 0.337$ ), i.e. those with two previous trials versus three. Simple slope analysis showed that the impact of previous adequate antidepressant trials on the decrease in MADRS scores was larger with augmentation than with switching.

We examined post hoc whether the number of previous adequate trials was a moderator of the change in MADRS scores,

expecting that the exposure–response relationship between these two variables would be stronger in the augmentation group compared with the switch group. Thus, we performed post hoc simple linear regression analysis between change in MADRS scores versus the number of previous adequate antidepressant trials in the augmentation and switch groups, including participants with two or more trials (Fig. 2). In the augmentation group, there was a significant association between the change in MADRS scores and the number of previous adequate trials ( $y = -1.5x + 10.9$ ;  $F_{1, 392} = 8.5$ ,  $P = 0.004$ , where  $y$  indicates the outcome variable (i.e. change in MADRS scores),  $x$  indicates the independent variable (i.e. number of previous adequate trials),  $F$  indicates the  $F$ -statistic, which is used to derive the significance of the regression model, and ‘1, 392’ following  $F$  indicates the degrees of freedom). In the switch group, there was no association between change in MADRS scores and the number of previous adequate trials ( $y = -0.7x + 6.7$ ;  $F_{1, 201} = 1.1$ ,  $P = 0.301$ ).

We also performed post hoc moderation analysis to examine whether the moderating effect on the different treatment strategies was driven by the different medications used in each treatment strategy (i.e. aripiprazole, bupropion and lithium were used for augmentation versus bupropion and nortriptyline used for switching). There was no moderating effect of the number of previous adequate antidepressant trials on the relationship between the medication used (aripiprazole versus bupropion) and change in MADRS scores in step 1 ( $b = -0.2$ ,  $t = -0.2$ ,  $P = 0.839$ , 95% CI  $[-1.9, 1.6]$ ). In step 2, lithium and nortriptyline did not significantly differ in regard to their effect on MADRS scores, negating the utility of moderation analysis.





**Fig. 2** Simple linear regression analysis of change in Montgomery–Åsberg Depression Rating Scale (MADRS) scores versus number of previous adequate antidepressant trials in the augmentation group ( $n=415$ ) and the switch group ( $n=212$ ). Error bars represent standard deviation. a. In the augmentation group, change in MADRS scores was significantly associated with the number of previous adequate trials ( $y = -1.5x + 10.9$ ;  $F_{1, 392} = 8.5$ ,  $P = 0.004$ ). In the switch group, change in MADRS scores was not associated with the number of previous adequate trials ( $y = -0.7x + 6.7$ ;  $F_{1, 201} = 1.1$ ,  $P = 0.301$ ), where  $y$  indicates the outcome variable (i.e. change in MADRS scores),  $x$  indicates the independent variable (i.e. number of previous adequate trials),  $F$  indicates the  $F$ -statistic, which is used to derive the significance of the regression model, and ‘1, 392’ following  $F$  indicates the degrees of freedom.

Previous adequate antidepressant trials are classified into augmentation or monotherapy. A further post hoc moderation analysis was performed to examine whether there was a moderating effect of a previous antidepressant trial being an augmentation trial or not. Of the 742 participants, 75 (10.1%) had at least one previous augmentation trial prior to enrolling, with 8 (1.1%) having two previous augmentation trials. There was no moderating effect of whether a participant had a past augmentation trial on the relationship between treatment strategies used and change in MADRS scores ( $b = -2.7$ ,  $t = -1.2$ ,  $P = 0.214$ , 95% CI  $[-7.1, 1.6]$ ).

Age ( $b = 0.2$ ,  $t = 1.4$ ,  $P = 0.161$ , 95% CI  $[-0.1, 0.4]$ ), executive dysfunction ( $b = 1.5$ ,  $t = 0.9$ ,  $P = 0.374$ , 95% CI  $[-1.8, 4.7]$ ), comorbid medical burden ( $b = 0.1$ ,  $t = 0.7$ ,  $P = 0.497$ , 95% CI  $[-0.2, 0.4]$ ) and comorbid anxiety ( $b = -0.40$ ,  $t = -0.24$ ,  $P = 0.810$ , 95% CI  $[-3.67, 2.87]$ ) did not have a moderating effect on the relationship between treatment strategy and change in MADRS scores.

Similarly, none of these five characteristics had a moderating effect on the relationship between treatment strategy and remission rates (age:  $b = 0.1$ ,  $z = 1.5$ ,  $P = 0.123$ , 95% CI  $[-0.02, 0.1]$ ; executive dysfunction:  $b = 0.3$ ,  $z = 0.7$ ,  $P = 0.485$ , 95% CI  $[-0.6, 1.3]$ ; comorbid medical burden:  $b = 0.1$ ,  $z = 1.4$ ,  $P = 0.153$ , 95% CI  $[-0.02, 0.2]$ ; comorbid anxiety:  $b = -0.42$ ,  $z = -0.87$ ,  $P = 0.38$ , 95% CI  $[-1.38, 0.53]$ ; and degree of treatment resistance:  $b = -0.26$ ,  $z = -1.10$ ,  $P = 0.27$ , 95% CI  $[-0.72, 0.20]$ ).

Sensitivity analyses showed that being of very old age (75 years or older) also did not moderate the relationship between treatment strategy and change in MADRS scores ( $b = 1.1$ ,  $t = 0.5$ ,  $P = 0.615$ , 95% CI  $[-3.2, 5.3]$ ) or remission ( $b = 0.7$ ,  $z = 1.1$ ,  $P = 0.281$ , 95% CI  $[-0.6, 1.9]$ ). Similarly, very impaired executive function (NIH Toolbox Fluid Cognition composite score 70 or lower) did not moderate the relationship between treatment strategy and change in MADRS scores ( $b = 0.01$ ,  $t = 0.003$ ,  $P = 0.997$ , 95% CI  $[-5.8, 5.8]$ ) or remission ( $b = 1.2$ ,  $z = 1.3$ ,  $P = 0.201$ , 95% CI  $[-0.6, 3.0]$ ).

## Discussion

We analysed data from the OPTIMUM randomised clinical trial comparing augmentation versus switching strategies in older adults

with TRD. We assessed five hypothesised moderators (age, executive dysfunction, comorbid medical burden, comorbid anxiety and degree of treatment resistance) on the relationship between the two strategies and treatment outcomes, as measured by change in MADRS scores or remission at week 10. Only the number of previous adequate antidepressant trials (i.e. degree of treatment resistance) was a significant moderator, where higher number of trials decreased one's likelihood of benefitting from augmentation over switching. Specifically, our findings suggest that the benefit of augmentation over switching is diminished in patients with three or more previous adequate antidepressant trials. We did not find a moderating effect of age, executive dysfunction, comorbid medical burden or comorbid anxiety, suggesting that these patient-level characteristics are less likely to impact whether someone will benefit from augmentation over switching.

Previous studies have reported that having had an adequate antidepressant trial predicts poor subsequent treatment outcomes in older adults.<sup>10,11</sup> Our results suggest that the number of previous antidepressant trials is also an important consideration, because those with fewer than three may benefit more from augmenting rather than switching. A meta-analysis reported the superiority of antipsychotic augmentation compared with ongoing antidepressant monotherapy, with an increasing number of up to four previous antidepressant trials.<sup>29</sup> Our analysis shows that, compared with switching, the effect of augmentation on symptom improvement decreases with increasing number of antidepressant trials (up to eight). This moderating effect appears to be specific to the treatment strategy and is not limited to the specific medications used. It also appears to be independent of whether a previous antidepressant trial was an augmentation trial or not. It has been consistently shown that treatment failure begets more treatment failures in depression.<sup>1,11,30</sup> This may be due to psychological factors, such as worsening treatment expectations,<sup>31</sup> or biological factors, such as progressive neurobiological changes or more severe immune alterations in those with TRD versus those without.<sup>32,33</sup> In a Cochrane review, augmentation strategies showed modest benefit compared with switching in patients with TRD.<sup>4</sup> Our results suggest that this modest benefit may disappear with a higher degree of treatment resistance.

We did not identify any moderators for remission. Because less than a third of participants attained remission, optimising the reduction of symptom severity by selecting the optimal treatment strategy is an important clinical goal in older adults with TRD. Our findings also indicate that augmentation and switching strategies perform both similarly and poorly in patients with multiple previous adequate antidepressant trials. This suggests that older patients with TRD with more than three adequate antidepressant trials should be considered for other intervention strategies, such as ketamine<sup>34</sup> or brain stimulation,<sup>35</sup> recognising that treatment response rates with these strategies decrease with treatment resistance as well.

Previous studies have reported age and executive dysfunction as moderators of treatment outcome to aripiprazole augmentation.<sup>6,7</sup> We did not find this using two different methods to examine age (either as a continuous variable or setting a threshold for ‘very old’), and using two different thresholds for executive dysfunction (impaired or very impaired). One possible explanation is that our participants were required to have had at least two failed previous antidepressant trials to be included in the study, while those other studies required only one previous trial. This may have resulted in the lack of effect of previously identified moderators in our more treatment-resistant sample. Furthermore, the study reporting age as a moderator examined a wider range of ages, from <50 to >65 years;<sup>6</sup> our sample consisted only of older adults with a smaller age range. Our previous study examined executive dysfunction as a





moderator of treatment outcome in late-life depression used the Trail Making test;<sup>7</sup> in that study we used the Fluid cognition score from the NIH Toolbox, which is a broad measure of fluid cognitive ability that allows for the separation of those with executive dysfunction and those without.<sup>16</sup> A meta-analysis showed that planning and organisation, but not cognitive flexibility, were associated with poor antidepressant treatment outcome in late-life depression,<sup>36</sup> suggesting that different domains of executive dysfunction may vary in their ability to act as moderators. Our previous studies have also reported comorbid physical illness and anxiety as negative predictors of treatment outcome in late-life depression.<sup>8,9,37</sup> While a predictor can also be a moderator, that is not necessarily the case: it depends on the sample and relationships being examined.<sup>38</sup> Our sample also consisted of older out-patients, typically with mild to moderate burden of comorbid physical illness. Having an unstable physical illness (e.g. an unmanaged cardiovascular condition) was also an exclusion criterion, which may have impacted our findings. Future studies are needed to validate these findings in older adults with a higher burden of physical illness. Notwithstanding the need for further exploration, our results suggest that, in older adults with TRD without severe executive dysfunction or medical comorbidities, neither their age, executive dysfunction, physical impairment nor level of anxiety pose additional barriers to benefiting from augmentation pharmacotherapy.

### Strengths and limitations

Our analysis has some strengths and limitations. The OPTIMUM study is the largest clinical trial to date in older adults with TRD.<sup>3</sup> It was a randomised clinical trial with standardised treatments focusing on broad strategies of augmentation versus switch, as opposed to specific medications, consistent with a growing consensus in the TRD field.<sup>4,5,39</sup> To our knowledge, this analysis is also the first to compare the effect of augmentation versus switch on depression symptom improvement in relation to the number of previous antidepressant trials. This preplanned analysis also had *a priori* focus on a predetermined list of five potential moderators selected based on previous studies reporting their association with treatment outcomes. Limitations include the retrospective collection of medication history, although collateral sources of information were used (e.g. health or pharmacy records) when available. Only antidepressants prescribed for up to 2 years were considered when rating the ATHF, which is a potential limitation given that some participants had major depressive episodes that lasted longer than 2 years. Furthermore, the version of the ATHF we used does not capture why medications were discontinued (e.g. lack of tolerability or lack of efficacy). Future studies should explore whether reasons for discontinuing antidepressant medications moderate the outcomes of future treatment. We used the same MADRS threshold of 10 for remission as the primary OPTIMUM analysis, which is a potential limitation because it may not indicate an absence of depression. In addition, 15.1% of participants did not have week 10 MADRS scores due to some having dropped out.<sup>3</sup> While the proportion of missing data was similar between switch and augmentation strategies, we cannot rule out the possibility of bias being introduced due to missing data. For example, it is reasonable to assume that participants who dropped out prior to week 10 had less improvement in their symptoms, influencing the strength of the interaction between previous antidepressant trials and treatment strategies. While recent TRD studies have compared augmentation versus switching rather than focusing on specific antidepressants,<sup>39</sup> some moderators may have differential effects on specific antidepressants when their therapeutic or adverse effects are directly related to them. For example, one study reported that

mirtazapine's superiority over selective serotonin reuptake inhibitors was moderated by differences in sleep and appetite, which are two prominent adverse effects of mirtazapine.<sup>40</sup> While our analysis showed that the moderating effect of the number of previous antidepressant trials is not due to the specific medication used, future studies could also look at the effect of other putative moderators on specific medications. Finally, our sample had limited diversity compared with the US or Canadian population, and participants were treated with four agents. Replicating this analysis in more varied samples with different agents will be important.

In conclusion, the findings of the OPTIMUM study can inform physicians and future studies on the selection of optimal treatment strategies for older patients with TRD. These findings identified readily available factors (i.e. the number of previous adequate antidepressant trials) that can be used to personalise clinical decision-making. Specifically, older patients with TRD with fewer than three previous adequate antidepressant trials may benefit more from augmentation strategies than switching to a different medication. Future studies are needed to identify effective treatments in patients with more than three previous adequate antidepressant trials.

**Helena K. Kim** , MD, PhD, Department of Psychiatry, Temerty Faculty of Medicine, University of Toronto, Toronto, Canada; **Jordan F. Karp**, MD, Department of Psychiatry, University of Arizona, Tucson, USA; **Helen Lavretsky**, MD, MS, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, USA; **Daniel M. Blumberger** , MD, MSc, Department of Psychiatry, Temerty Faculty of Medicine, University of Toronto, Toronto, Canada; and Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Canada; **Patrick J. Brown**, PhD, Department of Geriatric Psychiatry, Program on Healthy Aging and Late Life Brain Disorders, New York State Psychiatric Institute, Columbia University Medical Center, New York, USA; **Alastair J. Flint** , MD, Department of Psychiatry, Temerty Faculty of Medicine, University of Toronto, Toronto, Canada; and Centre for Mental Health, University Health Network, Toronto, Canada; **Emily Lenard**, Department of Psychiatry, Healthy Mind Lab, School of Medicine, Washington University in St Louis, St Louis, USA; **J. Philip Miller**, PhD, Department of Psychiatry, Healthy Mind Lab, School of Medicine, Washington University in St Louis, St Louis, USA; and Institute of Informatics, Data Science and Biostatistics, School of Medicine, Washington University in St Louis, St Louis, USA; **Charles F. Reynolds**, MD, Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, USA; **Steven P. Roose**, MD, Department of Geriatric Psychiatry, Program on Healthy Aging and Late Life Brain Disorders, New York State Psychiatric Institute, Columbia University Medical Center, New York, USA; **Eric J. Lenze**, MD, Department of Psychiatry, Healthy Mind Lab, School of Medicine, Washington University in St Louis, St Louis, USA; **Benoit H. Mulsant** , MD MS, Department of Psychiatry, Temerty Faculty of Medicine, University of Toronto, Toronto, Canada; and Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Canada

**Correspondence:** Benoit H. Mulsant. Email: [benoit.mulsant@utoronto.ca](mailto:benoit.mulsant@utoronto.ca)

First received 23 Jan 2025, final revision 31 Mar 2025, accepted 9 Apr 2025

### Data availability

The data for this study are available upon request following submission and approval of a data request proposal. Please contact the corresponding author, B.H.M.

### Acknowledgements

We thank members of the Data and Safety Monitoring Board (DSMB): Madhukar Trivedi, M., UT Southwestern (Chair); Mim Schwartz, Pittsburgh (Community Stakeholder Member); J Craig Nelson, MD, University of California in San Francisco (UCSF); Jeff Williamson, MD, Wake Forest University; Warren Taylor, MD, Vanderbilt University; and Knashawn Morales, PhD, University of Pennsylvania.

### Author contributions

H.K.K.: Conceptualisation; data curation; formal analysis; methodology; software; validation; visualisation; writing – original draft, review and editing; approval of final draft. J.F.K.: investigation; conceptualisation; supervision; writing – review and editing; approval of final draft. H.L.: funding acquisition; investigation; conceptualisation; supervision; writing – review and editing; approval of final draft. D.M.B.: investigation; writing – review and editing; approval of final draft. P.J.B.: investigation; writing – review and editing; approval of final draft. A.J.F.: investigation; writing – review and editing; approval of final draft. E. Lenard: data curation; project administration; resources; writing – review and editing; approval of final draft. J.P.M.: data curation; formal analysis; funding acquisition; methodology; writing – review and editing; approval of final draft. C.F.R.: funding acquisition; investigation; supervision; writing – review and editing; approval of final draft. S.P.R.: funding acquisition; investigation; supervision;

writing – review and editing; approval of final draft. E. Lenze: funding acquisition; investigation; supervision; writing – review and editing; approval of final draft. B.H.M.: conceptualisation; funding acquisition; investigation; methodology; supervision; validation; visualisation; writing – original draft, review and editing; approval of final draft.

## Funding

The original clinical trial was supported by a Patient-Centered Outcomes Research Institute (PCORI) award (no. TRD-1511-33321). This work was also supported in part by US National Institutes of Health (NIH) grant nos. MH114966, MH114969, MH114970, MH114980, MH114981, MH122686 and MH137353. H.K.K. was supported by a fellowship award from the Canadian Institutes of Health Research. B.H.M. was supported by the Labatt Family Chair in Biology of Depression in Late-Life Adults at the University of Toronto.

## Declaration of interest

The original clinical trial was supported by a PCORI award (no. TRD-1511-33321). This work was also supported in part by US NIH grant nos. MH114966, MH114969, MH114970, MH114980, MH114981, MH122686 and MH137353. D.M.B. reports grants from Canadian Institutes of Health Research (CIHR) and the Temerty family through the Centre for Addiction and Mental Health (CAMH) Foundation during the conduct of the study, non-financial support from Magventure (in-kind equipment support for investigator-initiated research), grants from Brainsway (principal investigator of an investigator-initiated study and site principal investigator for sponsored clinical trials), NIH, Brain Canada Foundation, Campbell Family Research Institute and PCORI outside the submitted work. D.M.B. also received medication supplies for an investigator-initiated trial from Indivior and participated in advisory boards for Janssen and Welcony. D.M.B. received additional support from the National Institute of Mental Health OPTIMUM NEURO grant (no. R01MH114980). A.J.F. has received grant support from NIH, PCORI, CIHR, Brain Canada, Ontario Brain Institute and the Alzheimer's Association. J.F.K. reports scientific advising for J&J Innovative Neuroscience (formerly Janssen) and Otsuka. J.F.K. also receives research support from J&J Innovative Neuroscience (formerly Janssen) and receives compensation from *Journal of Clinical Psychiatry* and *American Journal of Geriatric Psychiatry* for editorial board service. H.K.K. reports grant support from the PSI foundation (R23-21) and is supported by CIHR and the Temerty Faculty of Medicine (Chisholm Memorial Fellowship). E. Lenze received additional support from the Taylor Family Institute for Innovative Psychiatric Research at Washington University School of Medicine, and from the Washington University Institute of Clinical and Translational Sciences (grant no. UL1TR002345) from the National Center for Advancing Translational Sciences of NIH. P.M. has no financial relationships with commercial interests. B.H.M. received additional support from the Labatt Family Chair in Biology of Depression in Late-Life Adults at the University of Toronto. B.H.M. also holds and receives support from the Labatt Family Chair in Biology of Depression in Late-Life Adults at the University of Toronto and currently receives, or has received during the past 3 years, research support from Brain Canada, the CAMH Foundation, CIHR, US NIH, Capital Solution Design LLC (software used in a study funded by CAMH Foundation) and HAPPYneuron (software used in a study funded by Brain Canada). H.L. received support from grant nos. K24 AT009198, R01 AT008383 and R01 MH114981 from US NIH. C.F.R. receives payment from the American Association of Geriatric Psychiatry as Editor-in-Chief of the *American Journal of Geriatric Psychiatry* and royalty income for intellectual property as co-inventor of the Pittsburgh Sleep Quality Index. The other authors have no conflicts of interest to disclose.

## Transparency declaration

The manuscript is an accurate and transparent account of the reported study. No important aspects of the study were omitted. Discrepancies from the registered analysis plan have been reported. Clinical Trial Registration: clinicaltrials.gov NCT02960763 (<https://clinicaltrials.gov/study/NCT02960763>).

## References

- McIntyre RS, Alsuwaidan M, Baune BT, Berk M, Demyttenaere K, Goldberg JF, et al. Treatment-resistant depression: definition, prevalence, detection, management, and investigational interventions. *World Psychiatry* 2023; **22**: 394–412.
- Cristancho P, Lenard E, Lenze EJ, Miller JP, Brown PJ, Roose SP, et al. Optimizing outcomes of treatment-resistant depression in older adults (OPTIMUM): study design and treatment characteristics of the first 396 Participants randomized. *Am J Geriatr Psychiatry* 2019; **27**: 1138–52.
- Lenze EJ, Mulsant BH, Roose SP, Lavretsky H, Reynolds CF, Blumberger DM. Antidepressant augmentation versus switch in treatment-resistant geriatric depression. *N Engl J Med* 2023; **388**: 1067–79.
- Davies P, Ijaz S, Williams CJ, Kessler D, Lewis G, Wiles N. Pharmacological interventions for treatment-resistant depression in adults. *Cochrane Database Syst Rev* 2019; **12**: CD010557.
- Voineskos D, Daskalakis ZJ, Blumberger DM. Management of treatment-resistant depression: challenges and strategies. *Neuropsychiatr Dis Treat* 2020; **16**: 221–34.
- Zisook S, Johnson GR, Tal I, Hicks P, Chen P, Davis L, et al. General predictors and moderators of depression remission: a VAST-D report. *Am J Psychiatry* 2019; **176**: 348–57.
- Kaneriya SH, Robbins-Welty GA, Smagula SF, Karp JF, Butters MA, Lenze EJ, et al. Predictors and moderators of remission with aripiprazole augmentation in treatment-resistant late-life depression: an analysis of the IRL-GREY randomized clinical trial. *JAMA Psychiatry* 2016; **73**: 329–36.
- Reynolds CF, Dew MA, Pollock BG, Mulsant BH, Frank E, Miller MD. Maintenance treatment of major depression in old age. *N Engl J Med* 2006; **354**: 1130–8.
- Kim HK, Blumberger DM, Fitzgerald PB, Mulsant BH, Daskalakis ZJ. Antidepressant treatment outcomes in patients with and without comorbid physical or psychiatric disorders: a systematic review and meta-analysis. *J Affect Disord* 2021; **295**: 225–34.
- Kim HK, Blumberger DM, Karp JF, Lenze E, Reynolds CF, Mulsant BH. Venlafaxine XR treatment for older patients with major depressive disorder: decision trees for when to change treatment. *Evid Based Ment Health* 2022; **25**: 156–62.
- Buchalter ELF, Oughli HA, Lenze EJ, Dixon D, Miller JP, Blumberger DM, et al. Predicting remission in late-life major depression: a clinical algorithm based upon past treatment history. *J Clin Psychiatry* 2019; **80**: 1499.
- Coley RY, Boggs JM, Beck A, Hartzler AL, Simon GE. Defining success in measurement-based care for depression: a comparison of common metrics. *Psychiatr Serv* 2020; **71**: 312–8.
- Kim HK, Zai G, Muller DJ, Husain MI, Lam RW, Frey BN, et al. Identification of endocannabinoid predictors of treatment outcomes in major depressive disorder: a secondary analysis of the first Canadian biomarker integration network in depression (CAN-BIND 1) study. *Pharmacopsychiatry* 2022; **55**: 297–303.
- Papakostas GI, Trivedi MH, Shelton RC, Iosifescu DV, Thase ME, Jha MK, et al. Comparative effectiveness research trial for antidepressant incomplete and non-responders with treatment resistant depression (ASCERTAIN-TRD) a randomized clinical trial. *Mol Psychiatry* 2024; **29**: 2287–95.
- Uher R, Perlis RH, Placentino A, Dernovsek MZ, Henigsberg N, Mors O, et al. Self-report and clinician-rated measures of depression severity: can one replace the other? *Depress Anxiety* 2012; **29**: 1043–9.
- Heaton RK, Akshoomoff N, Tulsky D, Mungas D, Weintraub S, Dikmen S, et al. Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. *J Int Neuropsychol Soc* 2014; **20**: 588–98.
- Miller MD, Paradis CF, Houck PR, Mazumdar S, Stack JA, Rifai AH, et al. Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. *Psychiatry Res* 1992; **41**: 237–48.
- Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol* 2010; **63**: 1179–94.
- Dew RE, Kramer SI, McCall WV. Adequacy of antidepressant treatment by psychiatric residents: the antidepressant treatment history form as a possible assessment tool. *Acad Psychiatry* 2005; **29**: 283–8.
- Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; **134**: 382–9.
- Hawley CJ, Gale TM, Sivakumaran T. Defining remission by cut off score on the MADRS: selecting the optimal value. *J Affect Disord* 2002; **72**: 177–84.
- Little R, Schenker N. *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. Plenum Press, 1995.
- McClintock S, Ma ZQ, Rieger RH. Incorrect inference in prevalence trend analysis due to misuse of the odds ratio. *Ann Epidemiol* 2016; **26**: 136–40.
- Hayes AF. *Introduction to Mediation, Moderation, and Conditional Process Analysis. A Regression-Based Approach* 3rd ed. Guilford Press, 2022.
- Zhou J, Yang Y, Qiu X, Yang X, Pan H, Ban B, et al. Relationship between anxiety and burnout among chinese physicians: a moderated mediation model. *PLoS One* 2016; **11**: e0157013.
- Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990; **1**: 43–6.
- Duru G, Fantino B. The clinical relevance of changes in the Montgomery-Asberg Depression Rating Scale using the minimum clinically important difference approach. *Curr Med Res Opin* 2008; **24**: 1329–35.
- Hayes AF, Matthes J. Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behav Res Methods* 2009; **41**: 924–36.
- Wang HR, Woo YS, Ahn HS, Ahn IM, Kim HJ, Bahk WM. Can atypical antipsychotic augmentation reduce subsequent treatment failure more effectively among depressed patients with a higher degree of treatment resistance? A meta-analysis of randomized controlled trials. *Int J Neuropsychopharmacol* 2015; **18**: pyv023.
- Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one

- or several treatment steps: a STAR\*D report. *Am J Psychiatry* 2006; **163**: 1905–17.
- 31 Mollica A, Ng E, Burke MJ, Nestor SM, Lee H, Rabin JS, et al. Treatment expectations and clinical outcomes following repetitive transcranial magnetic stimulation for treatment-resistant depression. *Brain Stimul* 2024; **17**: 752–9.
  - 32 Chamberlain SR, Cavanagh J, de Boer P, Mondelli V, Jones DNC, Drevets WC, et al. Treatment-resistant depression and peripheral C-reactive protein. *Br J Psychiatry* 2019; **214**: 11–9.
  - 33 Howes OD, Thase ME, Pillinger T. Treatment resistance in psychiatry: state of the art and new directions. *Mol Psychiatry* 2022; **27**: 58–72.
  - 34 Oughli HA, Gebara MA, Ciarleglio A, Lavretsky H, Brown PJ, Flint AJ, et al. Intravenous ketamine for late-life treatment-resistant depression: a pilot study of tolerability, safety, clinical benefits, and effect on cognition. *Am J Geriatr Psychiatry* 2023; **31**: 210–21.
  - 35 Blumberger DM, Mulsant BH, Thorpe KE, McClintock SM, Konstantinou GN, Lee HH, et al. Effectiveness of standard sequential bilateral repetitive transcranial magnetic stimulation vs bilateral theta burst stimulation in older adults with depression: the FOUR-D randomized noninferiority clinical trial. *JAMA Psychiatry* 2022; **79**: 1065–73.
  - 36 Pimontel MA, Rindskopf D, Rutherford BR, Brown PJ, Roose SP, Sneed JR. A meta-analysis of executive dysfunction and antidepressant treatment response in late-life depression. *Am J Geriatr Psychiatry* 2016; **24**: 31–41.
  - 37 Andreescu C, Lenze EJ, Dew MA, Begley AE, Mulsant BH, Dombrovski AY, et al. Effect of comorbid anxiety on treatment response and relapse risk in late-life depression: controlled study. *Br J Psychiatry* 2007; **190**: 344–9.
  - 38 Courtney DB, Watson P, Krause KR, Chan BWC, Bennett K, Gunlicks-Stoessel M, et al. Predictors, moderators, and mediators associated with treatment outcome in randomized clinical trials among adolescents with depression: a scoping review. *JAMA Netw Open* 2022; **5**: e2146331.
  - 39 Ruberto VL, Jha MK, Murrough JW. Pharmacological treatments for patients with treatment-resistant depression. *Pharmaceuticals (Basel)* 2020; **13**: 116.
  - 40 Hieronymus F, Lisinski A, Eriksson E. Impact of sedative and appetite-increasing properties on the apparent antidepressant efficacy of mirtazapine, selective serotonin reuptake inhibitors and amitriptyline: an item-based, patient-level meta-analysis. *eClinicalMedicine* 2024; **77**: 102904.