Am J of Geriatric Psychiatry 33:4 (2025) 361–371

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ajgponline.org

Regular Research Article

# The Cognitive Profile of Older Adults With Treatment-Resistant Depression: An Analysis of the OPTIMUM Randomized Controlled Trial

***Nicholas J. Ainsworth, MD, Hanadi Oughli, MD, Helen Lavretsky, MD, MS, Daniel M. Blumberger, MD, MSc, Patrick J. Brown, PhD, Meryl A. Butters, PhD, Jordan F. Karp, MD, Emily Lenard, MSW, Eric J. Lenze, MD, Mary Pat McAndrews, PhD, J. Philip Miller, AB, Bruce G. Pollock, MD, PhD, Charles F. Reynolds, MD, Benoit H. Mulsant, MD, MS, on behalf of the OPTIMUM Research Group***

**ARTICLE INFO**

*Article history:*
Received April, 17 2024
Revised September, 14 2024
Accepted September, 23 2024

*Key Words:*
Aging
clinical trials
cognition
depression
executive function

**ABSTRACT**

***Objective:*** *Major depressive disorder in older adults (late-life depression; LLD) is frequently associated with cognitive impairment, and some deficits (e.g., executive function) have been associated with a higher level of treatment resistance. However, the cognitive profile of treatment-resistant LLD (TR-LLD) has not been characterized. We hypothesized that patients with TR-LLD would show deficits in cognitive function, especially executive function, and that executive function deficits would predict poorer response to pharmacotherapy.* ***Design:*** *Secondary analysis of baseline cognitive data from OPTIMUM, a multicenter RCT evaluating pharmacotherapy strategies for TR-LLD.* ***Setting:*** *Five outpatient academic medical centers (4 US, 1 Canada).* ***Participants:*** *About 369 participants aged 60 and older from the OPTIMUM study.* ***Measurements:*** *Baseline scores on individual tasks and composite scores from the NIH Toolbox-Cognition Battery were transformed into demographically-adjusted T-scores and compared to*

From the Centre for Addiction and Mental Health (NJA. DMB, BGP, BHM), Toronto, ON; Department of Psychiatry (NJA, DMB, BGP, BHM), University of Toronto, Toronto, ON; Department of Psychiatry and Biobehavioral Sciences (HO, HL), University of California, Los Angeles, CA; Department of Psychiatry, Columbia University College of Physicians and Surgeons (PJB), New York, NY; Department of Psychiatry (MAB, CFR), University of Pittsburgh School of Medicine, Pittsburgh, PA; Department of Psychiatry (JFK), College of Medicine, University of Arizona, Tucson, AZ; Department of Psychiatry (EL, EJL), Washington University School of Medicine, St. Louis, MO; Department of Psychology (MPM), University of Toronto, Toronto, ON; and the Institute for Informatics, Data Science and Biostatistics (JPM), Washington University School of Medicine, St. Louis, MO. Send correspondence and reprint requests to Nicholas J. Ainsworth, MD, Department of Psychiatry, Mount Saint Joseph Hospital, 3080 Prince Edward Street, Vancouver, BC V5T 3N4. e-mail: nickjainsworth@alumni.ubc.ca

© 2024 The Authors. Published by Elsevier Inc. on behalf of American Association for Geriatric Psychiatry. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

https://doi.org/10.1016/j.jagp.2024.09.018

*published norms. Impairments in the set shifting and inhibitory control tasks were investigated as predictors of depressive symptom change following treatment using ANCOVA models.* ***Results:*** *Participants had low performance on tasks evaluating inhibitory control, processing speed, verbal/nonverbal memory, and the fluid composite, but normative performance on working memory and set shifting. Participants had high estimated premorbid IQ (superior Performance on oral reading recognition). Age and physical comorbidity negatively associated with processing speed. Impairments in set shifting predicted less improvement in depressive symptoms; impairments in inhibitory control did not.* ***Conclusions:*** *Participants with TR-LLD presented with broad cognitive deficits relative to healthy norms. Given poorer outcomes following standard pharmacotherapy associated with impaired set shifting, future research needs to identify alternative treatment strategies.* (Am J Geriatr Psychiatry 2025; 33:361−371)

**Highlights**

- **What is the primary question addressed by this study?**
  What is the impact of treatment-resistant late-life depression (TR-LLD) on cognitive function?
- **What is the main finding of this study?**
  Individuals with TR-LLD exhibit broad-based cognitive deficits, including in some areas of executive function but not others.
- **What is the meaning of the finding?**
  Cognitive deficits in TR-LLD present a common clinical challenge that may impact on clinical course.

## INTRODUCTION

Major depressive disorder in older adults (late-life depression; LLD) is a risk factor for cognitive decline and dementia.[1] Many patients with LLD present with some cognitive impairment, and up to 40% meet diagnostic criteria for Mild Cognitive Impairment (MCI), conferring additional risk for future dementia.[2] Accordingly, the cognitive function of individuals suffering from LLD is a subject of ongoing interest.

Previous studies have shown that multiple cognitive domains can be affected in LLD, including attention, information processing speed, learning and memory, and executive function.3 With respect to executive function, several studies have found that patients with LLD frequently present with specific impairments in inhibitory control [3,4] and set shifting.[3] Comorbid executive dysfunction has been associated with poorer response to treatment with antidepressant medication,[4,5] an effect possibly mediated by poor treatment adherence.[6] Furthermore, executive dysfunction frequently persists even with treatment of LLD.[7–9]

To date, there are limited published data on cognitive function in individuals with treatment-resistant LLD (TR-LLD). The cognitive profile of these patients is important to understand in part because treatment resistance is highly prevalent and may be associated with underlying neurodegenerative changes related to Alzheimer's Disease and/or vascular disease.[10] Thus, cognitive impairment in TR-LLD may be a both a marker of treatment resistance and target for treatment.

In this context, we characterized domains of cognitive function in participants in the Optimizing Outcomes of Treatment-Resistant Depression in Older Adults (OPTIMUM) study.[11] In this analysis, we aimed to: 1) describe the cognitive profile of participants with TR-LLD across multiple cognitive domains; and 2) investigate whether baseline cognitive function predicts change in depressive symptoms

following an acute course of pharmacotherapy. We tested the hypotheses that: 1a) participants with TR-LLD would exhibit impaired cognitive function at baseline, based on published norms; 1b) participants with TR-LLD would demonstrate greatest impairment in measures of executive function, including inhibitory control and set shifting; and 2) based on our previous findings linking executive dysfunction and treatment resistance in LLD[10] participants with TR-LLD and impaired performance in measures of executive function (i.e., inhibitory control and set shifting) at baseline would show less improvement in depressive symptoms following protocolized pharmacotherapy than those with TR-LLD and intact executive function. We also explored whether sociodemographic and clinical characteristics explain the associations between cognitive impairment and TR-LLD.

## METHODS

### OPTIMUM Study

The design and implementation of the OPTIMUM study have been described in detail elsewhere.[11,12] Briefly, this was a pragmatic, randomized controlled trial (RCT) conducted at 5 sites (4 in the US, 1 in Canada) that compared the effectiveness of augmentation and switch strategies for TR-LLD. Participants were excluded if they showed evidence of possible dementia (based on a Short Blessed Test score >10), had unstable medical illness, or had a diagnosis of Parkinson's disease. Step 1 compared aripiprazole augmentation, bupropion augmentation, or switch to bupropion; Step 2 compared lithium augmentation to nortriptyline switch. Participants who were ineligible for Step 1 due to a previous medication trial were enrolled in Step 2 directly, while participants who did not remit after Step 1 were offered subsequent enrolment in Step 2 as a continuation of their participation. Each step lasted 10 weeks and participants completed the Montgomery-Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg[13]) at baseline and after 10 weeks. At baseline, they also completed the Patient-Reported Outcomes Measurement Information System (PROMIS) anxiety questionnaire[14] and the Cumulative Illness Rating Scale—Geriatric (CIRS-G).[15] Participants in the OPTIMUM trial were individuals age 60 or older with a diagnosis of major depressive disorder and treatment-resistant depression, as defined by a lack of remission of major depression after two or more antidepressant trials of adequate dose and duration.

### Cognitive Measures

All participants were asked to complete the NIH Toolbox Cognition Battery, Version 2 (NIHTB-CB) at baseline, prior to beginning their initial treatment in either Step 1 or Step 2, depending on eligibility. Only Step 1 participants and those who enrolled directly in Step 2 were included in this analysis. Because, as has been previously reported,[16] this trial utilized remote procedures to improve accessibility for individuals for whom transportation to the study site was not feasible, many participants were unable to complete the NIHTB-CB due to the need for on-site attendance for this instrument. This instrument was selected because of its extensive validation including available normative data, including in older adults.[17] Participants were asked to complete six tasks: (1) the Dimensional Change Card Sort Test ("Card Sort task"; a set shifting measure of cognitive flexibility); (2) the Flanker Inhibitory Control and Attention Test ("Flanker task"; inhibitory control and selective attention); (3) the List Sorting Working Memory Test ("List Sorting task"; working memory); (4) the Pattern Comparison Processing Speed Test ("Pattern Comparison task"; processing speed); (5) the Picture Sequence Memory Test ("Picture Sequence task"; verbal and nonverbal episodic memory); and (6) the Oral Reading Recognition Test ("Oral Reading task"; an oral word recognition task as a surrogate of premorbid intellectual function. All cognitive testing was completed on a tablet. Version 2.1 was used for all tests. For the Picture Sequence task, three forms with different sets of stimuli were used (Forms A, B, and C) and rotated intraindividually between assessments to mitigate practice effects.

We collected baseline scores for the fluid composite performance and the 6 individual tasks of the NIH Toolbox Cognition Battery for all participants who completed these assessments. We used the fluid composite score computed by the NIH Toolbox software, which averages the normalized scores of each measure completed in a given testing session,[18] with the exception of the Oral Reading Recognition task. For

individual tasks, where individual component scores were available (e.g., the reaction time and accuracy scores for the Flanker task), the combination scale score was used. To characterize the sample at baseline, we used the T-scores derived from the test norms[19] for each task adjusted for age, self-reported gender, and education, with a mean of 50 and a standard deviation of 10.

### Statistical Analysis

Of 740 participants in the OPTIMUM study (619 from Step 1 and 121 from Step 2), we included in the analysis 369 participants (315 from Step 1 and 54 from Step 2) who completed cognitive testing at baseline. We compared mean T scores for each baseline test within our sample against population norms using single-sample Student's t tests against a standardized population mean of 50, using Bonferroni correction for multiple comparisons. Correlations of cognitive scores with sociodemographic and clinical variables of interest were screened using a univariate correlation matrix; variables that were correlated ($p < 0.10$) with each cognitive score were included in a final multivariate linear regression model (Pearson method), together with three variables of *a priori* interest: baseline MADRS, PROMIS, and CIRS-G scores.

Next, we compared the change in MADRS scores following treatment in participants with versus without significant impairment (defined as a T-score of <40, i.e., at least one standard deviation below the standardized mean) in performance on two tests of executive function: the Card Sort task (set shifting) and the Flanker Task (inhibitory control). For each test, we analyzed the main effect of significant impairment in baseline performance on the change in MADRS score from baseline to the end of the treatment phase, using ANCOVA models adjusted for age, self-reported gender, education, baseline PROMIS score, baseline MADRS score, and other sociodemographic or clinical variables if they were correlated with the MADRS score change at a p-value of <0.1.

## RESULTS

### Summary of Included Sample

Of the 369 participants included in this analysis, the majority were female (239; 64.8%), with a mean (SD) age of 69.1 (6.9) years and a mean (SD) 14.7 (2.9) years of formal education. Table 1 presents the baseline sociodemographic and clinical characteristics of the sample. These characteristics were broadly similar

**TABLE 1. Baseline Sociodemographic and Clinical Characteristics of OPTIMUM Cognitive Sample**

| Sociodemographic | Step 1 (n = 315) | Step 2 (n = 54) |
|---|---|---|
| Self-identified female gender, n (%) | 203 (64.4) | 36 (66.7) |
| Age in years, mean (SD) | 69.0 (6.4) | 67.5 (4.2) |
| Education in years, mean (SD) | 14.8 (3.1) | 15.5 (2.4) |
| Married, n (%) | 118 (37.4) | 18 (33.3) |
| Race, n (%) | | |
| Asian | 3 (1.0) | 0 (0.0) |
| Black | 24 (7.7) | 0 (0.0) |
| Hawaiian/Pacific Islander | 0 (0.0) | 0 (0.0) |
| Indigenous | 0 (0.0) | 0 (0.0) |
| Multirace | 6 (1.6) | 1 (1.9) |
| Other | 14 (3.9) | 0 (0.0) |
| White | 268 (85.6) | 52 (96.3) |
| Clinical | | |
| MDD age of onset in years, mean (SD) | 32.6 (20.4) | 24.7 (15.3) |
| Number of adequate medication trials, mean (SD) | 2.4 (0.9) | 2.5 (1.0) |
| CIRS-G score, mean (SD)[a] | 8.3 (4.8) | 7.6 (4.9) |
| PROMIS anxiety questionnaire T-score, mean (SD) | 64.2 (6.9) | 64.0 (7.3) |
| MADRS score at baseline, mean (SD)[a] | 22.6 (7.3) | 22.8 (7.5) |
| MADRS score at end of treatment, mean (SD) | 16.2 (9.0) | 17.6 (9.0) |

SD: standard deviation; MDD: major depressive disorder; CIRS-G: Cumulative Illness Rating Scale – Geriatric; PROMIS, patient-reported outcomes measurement information system; MADRS, Montgomery-Åsberg depression rating scale.

[a] Significant difference between participants with and without baseline cognitive data.

**TABLE 2. Comparison of Baseline Scores on NIH Toolbox Cognitive Battery Tasks Against Population Norms**

| Cognitive Test | Test Statistic | df | p-value | Mean Difference | Effect Size (Cohen's d) | Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower | Upper |
| Card sorting task | t = 0.39 | 365 | 0.699 | 0.21 | 0.02 | −0.08 | 0.12 |
| Flanker task | t = −25.04 | 364 | <0.001 | −9.36[a] | −1.31 | −1.45 | −1.17 |
| Fluid composite | t = −10.75 | 288 | <0.001 | −6.29[a] | −0.63 | −0.76 | −0.51 |
| List sorting task | t = −2.54 | 360 | 0.012 | −1.2 | −0.13 | −0.24 | −0.03 |
| Pattern comparison task | t = −12.08 | 363 | <0.001 | −8.67[a] | −0.63 | −0.75 | −0.52 |
| Picture sequence task | t = −5.68 | 350 | <0.001 | −2.94[a] | −0.30 | −0.41 | −0.20 |

Notes: Test statistics are for single-sample t-tests using a standardized population mean T-score of 50 and hypothesis: ≠ 50.
Confidence intervals are for Cohen's d.
df, Degrees of freedom (differing values reflect different numbers of participants who completed each task).
Negative mean difference values correspond to a lower score on the task versus the standardized population mean.
[a] Significant difference compared with standardized population mean for each test, based on an alpha threshold (Bonferroni) of 0.0071.

between the 369 participants included in this analysis and the 371 participants who were not included because of missing cognitive data, except that included participants had a slightly lower level of physical comorbidity (CIRS-G mean [SD]: 8.2 [4.8] versus 9.1 [4.6], t = 2.58, df = 733, p = 0.01) and lower baseline MADRS score (mean [SD]: 22.6 [7.4] versus 23.7 [7.1], t = 2.08, df = 714, p = 0.04). Similarly, participants who entered the study at Step 1 or directly at Step 2 had minimal differences in sociodemographic and clinical characteristics as well as baseline cognitive scores: the only statistically significant difference across 15 comparisons was in years of education, which was slightly higher in participants directly entering Step 2 (mean [SD]: 15.2 [2.6] years versus 14.6 [2.9] years, t = 2.14, df = 725, p = 0.03).

## Baseline Cognitive Function

TR-LLD participants were significantly impaired on most cognitive domains at baseline, based on published demographic norms accounting for age, gender, and education (see Fig. 1A–F). A large effect size for impairment was seen in the Flanker task (inhibitory control) (mean [SD] T-score: 40.6 [7.1], t = −25.04, df = 364, p <0.001; Cohen's d = −1.31 [95% CI: −1.45, −1.17]), which remained significant after correction for multiple comparisons. Moderate effect sizes for impairment were also seen on the Pattern Comparison task (processing speed) (mean [SD] T-score: 41.3 [13.7], t = −12.08, df = 363, p <0.001; Cohen's d = −0.63 [95% CI: −0.75, −0.52]) and on the Fluid Composite (mean [SD] T-score: 43.7 [9.9], t = −10.75, df = 288, p <0.001; Cohen's d = −0.63 [95% CI: −0.76, −0.51]), with a small effect size on the Picture Sequence task (memory) (mean [SD] T-score: 47.1 [9.7], t = −5.68, df = 350, p < 0.001; Cohen's d = −0.30 [95% CI: −0.41, −0.20]); all of these remained significant after correction for multiple comparisons (See Table 2). No significant impairment was seen on the List Sort task (working memory) or Card Sort task (set shifting) after correction for multiple comparisons (see Table, Supplemental Digital Content 1 for further descriptive statistics). TR-LLD participants performed significantly better than average on the oral reading comprehension task (a surrogate for premorbid intellectual function) based on published norms (mean [SD] T-score: 57.9 [8.5], t = 17.66, df = 355, p <0.001; Cohen's d = 0.94 [95% CI: 0.81, 1.06]).

## Correlates of Baseline Cognitive Function in TR-LLD

Several sociodemographic and clinical variables were correlated with baseline scores on individual cognitive tasks following univariate linear regression (see Figure, Supplemental Digital Content 2, containing the univariate correlation matrix used to screen variables). However, only two of these associations remained significant after covariate adjustment: age (estimated coefficient [unstandardized weight]: −0.377 [95% CI: −0.620, −0.134], t = −3.05, df = 313, p = 0.002) and CIRS-G scores (estimated coefficient [unstandardized weight]: −0.473 [95% CI: −0.774, −0.172], t = −3.09, df = 313, p = 0.002) each were negatively associated with scores on the Pattern Comparison task.

**FIGURE 1. Distributions of cognitive scores for OPTIMUM participants at baseline. Each histogram shows the distribution of demographically-adjusted T-scores obtained in our sample from a task within the NIH Toolbox — Cognitive Battery: Fluid Composite (A), Card Sort task (B), Flanker task (C), List Sorting task (D), Pattern Comparison task (E), and Picture Sequence task (F). The standardized mean T-score for each task based on demographic norms is 50 (solid lines), with a standard deviation of 10. The mean T-score obtained in our sample for each task is indicated by a dashed line. ***Significant difference from standardized mean, p <0.001.**

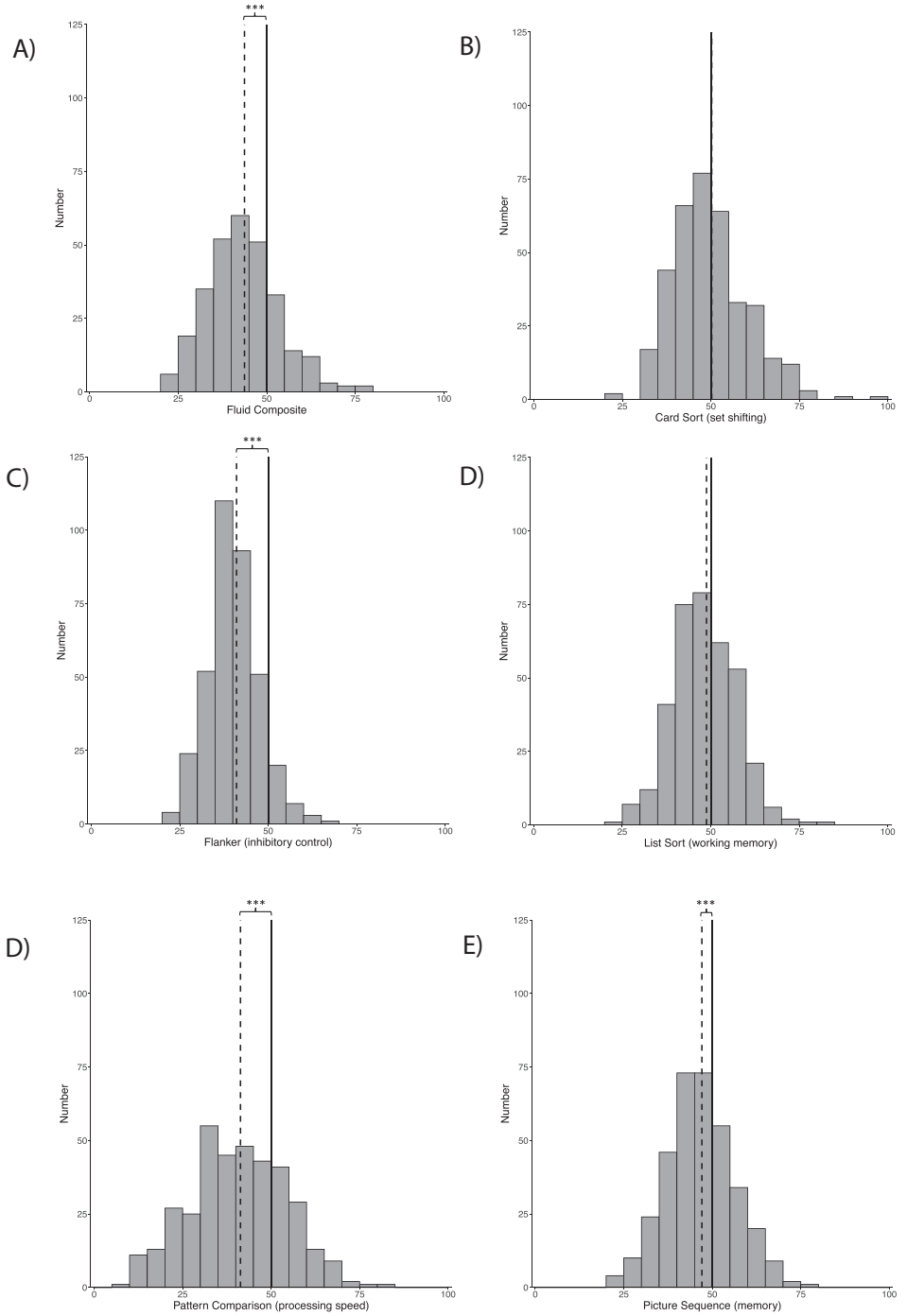
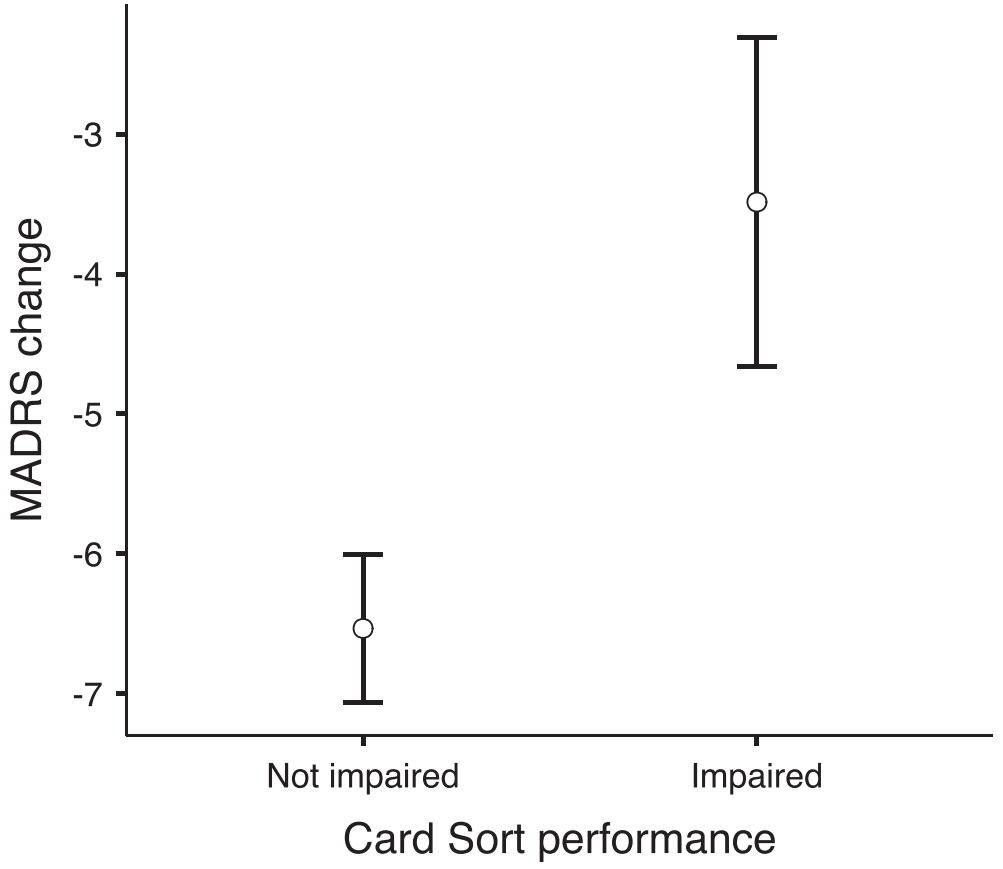


**FIGURE 2. Effect of baseline set-shifting impairment on change in MADRS score after treatment. Impaired −Card Sort task baseline T-score <40. Not Impaired −Card Sort task baseline T-score ≥40. Whiskers denote standard errors.**




### Does Executive Dysfunction Predict Less Depressive Symptom Improvement in TR-LLD?

Of 366 participants who completed the Card Sorting task, 58 (15.8%) showed impairment, as defined by a baseline T-score on the Card Sort task of <40, while 165 of 365 participants (45.2%) showed impairment on the Flanker task. Stratification of impairment by mild (T-score: 39−30), moderate (T-score: 29−20), or severe (T-score <20) revealed that most participants in the impaired range for each task showed mild impairment, with very few in the moderate impairment category and none in the severe category (see Table, Supplemental Digital Content 3).

Participants with impairment in set-shifting performance had less improvement in depressive symptoms following treatment than participants with better set-shifting performance (adjusted mean difference in MADRS score change = −3.05 points, df = 240, p = 0.02, Cohen's d = −0.41) (see Fig. 2). Conversely, participants with impairment in inhibitory control performance (T-score of <40 on the Flanker task) did not significantly differ in change in MADRS score compared with those with higher performance. These overall effects remained significant when the analysis was restricted to Step 1 participants only.

## DISCUSSION

We analyzed the cognitive profile of 369 participants with TR-LLD who were enrolled in the first

step of the OPTIMUM pharmacotherapy trial. Our analysis had two main findings. First, our TR-LLD sample exhibited impairment relative to healthy older adults in composite cognition and in specific domains based on test norms, with the largest degree of impairment seen in one area of executive function, inhibitory control (as measured by the Flanker Task). However, our sample did not exhibit impairment in another measure of executive function, set shifting (as measured by the Card Sort task). These impairments were also observed in individuals who were relatively well educated and demonstrated significantly superior performance on an oral word recognition task (reflecting premorbid general intelligence). In this treatment-resistant sample, we did not observe any consistent sociodemographic or clinical correlates of baseline cognitive function, except that older age and greater medical comorbidity were associated with worse information processing speed. Second, in our analysis of baseline executive function as a predictor of depressive symptom outcomes, baseline low performance in set-shifting (but not inhibitory control) predicted less improvement in depressive symptoms following pharmacotherapy for TR-LLD.

While the cognitive profile of patients with LLD has now been well characterized,[2,3] to our knowledge, this is the first description of the cognitive profile of patients with TR-LLD receiving pharmacotherapy. Prior work has characterized cognitive function in mid-aged adults with treatment resistant depression,[20,21] and the cognitive profile of a sample of older adults receiving ECT has also been reported.[22] Our analysis confirmed our hypothesis that participants with TR-LLD would demonstrate broad cognitive impairment at baseline, and we saw more widespread deficits than have typically been identified in non−treatment-resistant LLD.[2,3] However, we saw a divergence in tasks measuring components of executive function, with significant impairment seen with the Flanker Task (inhibitory control) but not with the Card Sort task (set-shifting). The impairment in inhibitory control in our sample is congruent with prior findings that older depressed patients with impaired inhibitory control are less likely to respond to initial antidepressant treatment,[4,5] which increases the likelihood that they would require further treatment as offered in the OPTIMUM trial. The divergence we saw in baseline performance on executive function tasks underlines the complexity of executive function and the importance of considering its various putative components. Based on our results, it would be inexact to describe our participants with TR-LLD as exhibiting "executive dysfunction" broadly, and more accurate to describe them as exhibiting an impairment in inhibitory control. Taken together, these findings suggest that individuals with TR-LLD experience specific and substantial impairments in cognitive function (including in inhibitory control). They also suggest that a high premorbid intelligence ("cognitive reserve"), as estimated here by the Oral Reading task, may not be enough to mitigate this deficit; however, it should be noted that this finding requires replication with other surrogate tests for cognitive reserve, which can estimate this differently[23,24] They also suggest that in individuals with TR-LLD, impairment is not seen exclusively in executive function; rather, cognitive impairment in this group appears to be broadly distributed among several cognitive domains, and some areas of executive function may be spared, particularly in individuals with higher premorbid intellectual functioning. This includes significant impairment in memory, the persistence of which despite adequate treatment has been previously described in the existing LLD literature, and which may suggest an increased risk of developing dementia.[8]

Our analysis of executive function in older patients as a predictor of response to TR-LLD pharmacotherapy yielded contrasting findings: participants with worse set-shifting experienced significantly less improvement in depressive symptoms following treatment than those without better set-shifting, while participants with worse inhibitory control experienced a similar degree of depressive symptom improvement as those with better inhibitory control. This is consistent with prior results from the IRL-GRey study[25] among participants who received aripiprazole augmentation following nonremission to a prospective trial of venlafaxine[10] However, it contrasts with results from several studies of patients with LLD without treatment resistance: most of these studies have found an association between specific impairments in inhibitory control and less depressive symptom improvement following pharmacotherapy and this has been confirmed in a meta-analysis.[5] Since inhibitory control predicts treatment resistance in

LDD, patients with TR-LLD may have a more consistent impairment in inhibitory control such that it precludes predicting treatment response in these patients; this may lead to inhibitory control having a more homogeneous impact on treatment response in TR-LLD versus non−treatment-resistant LLD. Another factor could contribute to the different relationship between cognitive function and depressive outcomes in patients with LLD or TR-LLD: both medications used as the first treatment step of OPTIMUM, (i.e., aripiprazole and bupropion) act on dopamine; thus, they may be better suited for treating depressive symptoms in those with impaired inhibitory control than selective serotonin reuptake inhibitors or other first-line antidepressants. Given the evidence supporting a relationship between dopaminergic transmission and inhibitory control,[26,27] future studies will need to determine whether dopaminergic medications may mitigate the deleterious impact on depressive symptom improvement attributed to executive dysfunction. Conversely, the worse outcomes experienced by the participants in our TR-LLD sample with poor set-shifting, consistent with a prior study,[10] suggest that this impairment may be a trait-like feature in TR-LLD that forms part of a highly treatment-resistant subgroup. It should be noted that this subgroup was relatively small in our study, suggesting that the adverse impact of this form of executive dysfunction may be relatively uncommon even in a TR-LLD sample.

Our results should be considered in the context of some limitations. First, complete NIH Toolbox Cognition Battery data were only available for a subset of OPTIMUM participants, limiting our sample size. However, comparison of participants who did and did not have these cognitive data revealed only small differences in baseline characteristics (0.6 years of education and 1.1 points on the MADRS). It is unlikely that these small differences would have any impact on our results; additionally, all our analyses controlled for potential clinical and sociodemographic confounders. Second, we did not have a sample of patients with non−treatment-resistant LLD to which we could directly compare the cognitive profile of our participants with TR-LLD; therefore, all analyses involved either participant scores based on test norms in healthy individuals,[19] or indirect reference to prior literature in non−treatment-resistant LLD samples. Third, we were limited in our ability to compare different levels of executive function impairment on treatment outcomes, as most participants were in the mild impairment range. Fourth, our measure of set-shifting, the Dimensional Change Card Sort Test, is a complex measure that may also capture some other cognitive abilities such as inhibition and processing speed[28−30]; however, our results are consistent with a prior study using a different measure of set-shifting (i.e., the Trail-Making Test Part B). Finally, our protocol utilized a single cognitive battery (the NIH Toolbox) and selected antidepressant agents, so these findings may not generalize to other samples utilizing different cognitive tests or medications.

## CONCLUSION

Cognitive impairment is a major clinical challenge in LLD. Our results confirm that cognitive impairment is substantial and widespread in patients with TR-LLD, with multiple domains affected, including some but not all measures of executive functioning. Our results also suggest that while poor baseline set-shifting predicts less improvement in depressive symptoms following pharmacotherapy for TR-LLD, poor inhibitory control may not decrease the likelihood of responding to medications that affect the dopaminergic system. In summary, our results provide a novel description and analysis of the cognitive profile of patients with TR-LLD, and they provide insights regarding management of these patients when they present with executive dysfunction.

## DISCLOSURES

*NJA receives support from the University of British Columbia Institute of Mental Health, CIHR-CANTRAIN, and Michael Smith Health Research BC. DMB receives research support from CIHR, NIH, Brain Canada and the Temerty Family through the CAMH Foundation and the Campbell Family Research Institute. He received research support and in-kind equipment support for an investigator-initiated study from Brainsway Ltd. He was the site principal investigator for three sponsor-initiated studies for Brainsway Ltd. He also received in-kind equipment support from Magventure for two investigator-initiated*

*studies. He received medication supplies for an investigator-initiated trial from Indivior. He is a scientific advisor for Sooma Medical. He is the Co-Chair of the Clinical Standards Committee of the Clinical TMS Society (unpaid). HL receives grant support from NIH and the US Department of Defense. EJL receives honoraria for consulting with Boehinger-Ingelheim, Merck, Pritikin ICR, Prodeo, IngenioRx; grant support from Janssen; and has a patent pending for Sigma1 agonists for COVID. BGP received research support from the Peter & Shelagh Godsoe Endowed Chair in Late-Life Mental Health, CAMH Foundation, and Discovery Fund, National Institute of Aging, Brain Canada, the Canadian Institutes of Health Research, the Alzheimer's Drug Discovery Foundation, the Ontario Brain Institute, the Centre for Aging and Brain Health Innovation, the Bright Focus Foundation, the Alzheimer's Society of Canada, the W. Garfield Weston Foundation, the Weston Brain Institute, the Canadian Consortium on Neurodegeneration in Aging and Genome Canada. Dr. Pollock receives honoraria from the American Geriatrics Society for book authorship and holds United States Provisional Patent Nos. 6/490,680, 17/396,030, and Canadian Provisional Patent No. 3,054,093 for a cell-based assay and kits for assessing serum anticholinergic activity. JFK has received research support from Janssen and served as an advisor to Biogen and Aifred Health. He receives compensation from Physician's Postgraduate Press and American Journal of Geriatric Psychiatry for editorial service. MPM receives research support from the Canadian Institutes of Health Research, The Weston Brain Institute, and the Krembil Research Institute. JPM receives grant support from NIH and PCORI. BHM holds and receives support from the Labatt Family Chair in Biology of Depression in Late-Life Adults at the University of Toronto. He currently receives or has received within the past 5 years research support from Brain Canada, the Canadian Institutes of Health Research, the CAMH Foundation, the Patient-Centered Outcomes Research Institute (PCORI), the US National Institute of Health (NIH), Capital Solution Design LLC (software used in a study founded by CAMH Foundation), and HAPPYneuron (software used in a study founded by Brain Canada). He has also been an unpaid consultant to Myriad Neuroscience.*

## AUTHOR CONTRIBUTIONS

*NJA contributed to study design, analysis, manuscript drafting and revision; HO contributed to study design, analysis, and manuscript revision; HL, DMB, PJB, JFK, EJL, MPM, BGP, and CFR contributed to study design and manuscript revision; EL contributed to data collection, analysis, and manuscript revision; JPM contributed to analysis and manuscript revision; BHM contributed to study design, manuscript drafting and revision.*

## DATA STATEMENT

*Due to confidentiality and funding agency constraints, these data cannot be made publicly available.*

## PREVIOUS PRESENTATION

*These data have not been previously presented orally or by poster at scientific meetings.*

## ACKNOWLEDGMENTS

*The authors would like to acknowledge the OPTIMUM trial team, in particular Michael Yingling and Dr. Torie Gettinger, for their assistance in data management.*

*This work was supported by a Patient-Centered Outcomes Research Institute (PCORI) Award (TRD-1511-33321).*

**Except for the first three authors and the last author, all the authors contributed equally and are listed in alphabetical order.*

## SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.jagp.2024.09.018.

## References

1. Livingston G, Huntley J, Sommerlad A, et al: Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020; 396(10248):413–446;doi:10.1016/S0140-6736(20)30367-6

2. Butters MA, Whyte EM, Nebes RD, et al: The nature and determinants of neuropsychological functioning in late-life depression. Arch Gen Psychiatry 2004; 61(6):587–595;doi:10.1001/archpsyc.61.6.587

3. Koenig AM, Bhalla RK, Butters MA: Cognitive functioning and late-life depression. J Int Neuropsychol Soc 2014; 20(5):461–467;doi:10.1017/S1355617714000198
4. Alexopoulos GS, Kiosses DN, Heo M, et al: Executive dysfunction and the course of geriatric depression. Biol Psychiatry 2005; 58(3):204–210;doi:10.1016/j.biopsych.2005.04.024
5. Pimontel MA, Rindskopf D, Rutherford BR, et al: A meta-analysis of executive dysfunction and antidepressant treatment response in late-life depression. Am J Geriatr Psychiatry 2016; 24(1):31–41;doi:10.1016/j.jagp.2015.05.010
6. Cristancho P, Lenze EJ, Dixon D, et al: Executive function predicts antidepressant treatment noncompletion in late-life depression. J Clin Psychiatry 2018; 79(3):e1–e7;doi:10.4088/JCP.16m11371
7. Aizenstein HJ, Butters MA, Wu M, et al: Altered functioning of the executive control circuit in late-life depression: episodic and persistent phenomena. Am J Geriatr Psychiatry 2009; 17(1):30–42
8. Bhalla RK, Butters MA, Mulsant BH, et al: Persistence of neuropsychologic deficits in the remitted state of late-life depression. Am J Geriatr Psychiatry 2006; 14(5):419–427;doi:10.1097/01.JGP.0000203130.45421.69
9. Ainsworth NJ, Marawi T, Maslej MM, et al: Cognitive outcomes after antidepressant pharmacotherapy for late-life depression: a systematic review and meta-analysis. Am J Psychiatry 2024; 181(3):234–245;doi:10.1176/appi.ajp.20230392
10. Kaneriya SH, Robbins-Welty GA, Smagula SF, 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(4):329–336; doi:10.1001/jamapsychiatry.2015.3447
11. Lenze EJ, Mulsant BH, Roose SP, et al: Antidepressant augmentation versus switch in treatment-resistant geriatric depression. New Eng J Med 2023; 388(12):1067–1079; doi:10.1056/nejmoa2204462
12. Cristancho P, Lenard E, Lenze EJ, 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(10):1138–1152; doi:10.1016/j.jagp.2019.04.005
13. Montgomery SA, Åsberg M: A new depression scale designed to be sensitive to change. Br J Psychiatry 1979; 134(4):382–389; doi:10.1192/bjp.134.4.382
14. Cella D, Riley W, Stone A, 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(11):1179–1194; doi:10.1016/j.jclinepi.2010.04.011
15. Miller MD, Paradis CF, Houck PR, et al: Rating chronic medical illness burden in geropsychiatric practice and research: application of the cumulative illness rating scale. Psychiatry Res 1992; 41(3):237–248;doi:10.1016/0165-1781(92)90005-N
16. Ainsworth NJ, Wright H, Tereshchenko K, et al: Recruiting for a randomized clinical trial for late-life depression during COVID-19: outcomes of provider referrals versus Facebook self-referrals. Am J Geriatr Psychiatry 2023; 31(5):366–371;doi:10.1016/j.jagp.2023.01.021
17. Weintraub S, Dikmen SS, Heaton RK, et al: Cognition assessment using the NIH Toolbox. Neurology 2013; 80(11);doi:10.1212/wnl.0b013e3182872ded, Suppl 3
18. Slotkin J, Nowinski CJ, Hays R, et al. NIH toolbox scoring and interpretation guide. 2012.
19. Casaletto KB, Umlauf A, Beaumont J, et al: Demographically corrected normative standards for the english version of the NIH toolbox cognition battery. J Int Neuropsychol Soc 2015; 21(5):378–391;doi:10.1017/S1355617715000351
20. McClintock SM, Cullum CM, Husain MM, et al: Evaluation of the effects of severe depression on global cognitive function and memory. CNS Spectr 2010; 15(5):304–313;doi:10.1017/S109285290002753X
21. Moreines JL, McClintock SM, Holtzheimer PE: Neuropsychologic effects of neuromodulation techniques for treatment-resistant depression: a review. Brain Stimul 2011; 4(1):17–27; doi:10.1016/j.brs.2010.01.005
22. Lisanby SH, McClintock SM, Alexopoulos G, et al: Neurocognitive effects of combined electroconvulsive therapy (ECT) and venlafaxine in geriatric depression: phase 1 of the PRIDE study. Am J Geriatr Psychiatry 2020; 28(3):304–316;doi:10.1016/j.jagp.2019.10.003
23. Manly JJ, Jacobs DM, Touradji P, et al: Reading level attenuates differences in neuropsychological test performance between African American and White elders. J Int Neuropsychol Soc 2002; 8(3):341–348;doi:10.1017/s1355617702813157
24. Jones RN, Manly J, Glymour MM, et al: Conceptual and measurement challenges in research on cognitive reserve. J Int Neuropsychol Soc 2011; 17(4):593–601;doi:10.1017/S1355617710001748
25. Lenze EJ, Mulsant BH, Blumberger DM, et al: Efficacy, safety, and tolerability of augmentation pharmacotherapy with aripiprazole for treatment-resistant depression in late life: a randomised, double-blind, placebo-controlled trial. Lancet 2015; 386(10011):2404–2412;doi:10.1016/S0140-6736(15)00308-6
26. Ott T, Nieder A: Dopamine and cognitive control in prefrontal cortex. Trends Cogn Sci 2019; 23(3):213–234;doi:10.1016/j.tics.2018.12.006
27. Parr AC, Calabro F, Tervo-Clemmens B, et al: Contributions of dopamine-related basal ganglia neurophysiology to the developmental effects of incentives on inhibitory control. Dev Cogn Neurosci 2022; 54(October 2021):101100;doi:10.1016/j.dcn.2022.101100
28. Ott LR, Schantell M, Willett MP, et al: Construct validity of the NIH toolbox cognitive domains: a comparison with conventional neuropsychological assessments. Neuropsychology 2022; 36(5):468–481;doi:10.1037/neu0000813
29. Zelazo PD, Anderson JE, Richler J, et al: NIH Toolbox Cognition Battery (CB): validation of executive function measures in adults. J Int Neuropsychol Soc 2014; 20(6):620–629;doi:10.1017/S1355617714000472
30. Scott EP, Sorrell A, Benitez A: Psychometric properties of the NIH toolbox cognition battery in healthy older adults: reliability, validity, and agreement with standard neuropsychological tests. J Int Neuropsychol Soc 2019; 25(8):857–867;doi:10.1017/S1355617719000614