Neurocognitive Correlates of Response to Treatment in Late-Life Depression

Tyler J. Story, Ph.D.1,2, Guy G. Potter, Ph.D.2, Deborah K. Attix, Ph.D.1,2, Kathleen A. Welsh-Bohmer, Ph.D.1,2,3, and David C. Steffens, M.D.4

1Division of Neurology, Duke University Medical Center.
2Division of Medical Psychology, Duke University Medical Center.
3Alzheimer’s Disease Research Center, Duke University Medical Center.
4Division of Geriatric Psychiatry, Duke University Medical Center.

Abstract

Depression is often associated with neurocognitive deficits in older adults, particularly in the domains of information processing speed, episodic memory, and executive functions. Greater neurocognitive dysfunction while depressed is associated with a less effective treatment response; however, questions remain about the specific variables that characterize patients showing low treatment response and persistent cognitive deficiencies.

OBJECTIVES—The authors examined neurocognitive variables that differentiated patients who showed robust versus weak responses to antidepressant therapy.

PARTICIPANTS—The baseline sample included 110 women and 67 men, with a mean age of 69.1 years (SD = 6.9) and mean education of 14 years (SD = 3.3).

DESIGN—Patients enrolled in a treatment study completed both a structured diagnostic assessment for depression and neuropsychological testing at study entry and one-year follow-up.

MEASUREMENTS—Clinicians rated patient depression using the MADRS. Neuropsychological assessments consisted of prose recall and percent retention (WMS-III Logical Memory), word-list recall, attention and visuomotor processing speed (Trail Making A, Symbol-Digit Modalities), and mental flexibility (Trail Making B).

INTERVENTIONS—Patients underwent treatment for depression following the guidelines of the Duke STAGED approach.

RESULTS—Individuals who demonstrated the greatest improvement in mood symptoms at follow-up exhibited better prose recall and faster processing speed at baseline than individuals who demonstrated weaker treatment responses. These differences remained after controlling for depression severity at both time-points.

CONCLUSION—The current results suggest that better pre-treatment cognitive function, particularly in verbal memory, is associated with a greater treatment response in late life depression.
Keywords
Depression; geriatric; cognitive decline; treatment response

Introduction
Optimizing recovery from depression is a topic of strong interest in geriatric psychiatry considering that adults with late-life depression (LLD) suffer higher rates of mortality and disability, and generate higher health care costs than older patients without mood disorders (1,2,3). One treatment challenge is that approximately one-third of older adults are resistant to typical pharmacological intervention (4). Treatment of LLD is complicated by age-related structural brain changes that may contribute to persistent depression. For example, LLD has been associated with ischemic pathology in prefrontal (5,6,7) and subcortical structures (8–15), along with vital communicating pathways in between. Greater severity of vascular pathology in these frontostriatal circuits is associated with slowed treatment response (16–18). There is also evidence that persisting depressive episodes are associated with hippocampal atrophy (19), possibly as a result of impaired regulation of glucocorticoid secretion (20–21). Because abnormalities of mood-critical structures in the frontostriatal and hippocampal regions appears to play a role in slowing or limiting an individual’s recovery from depression, it is important to identify parallel clinical characteristics of depressed individuals who may be at risk of poor treatment response.

Individuals with underlying frontostriatal and hippocampal dysfunction may be predisposed to poor treatment response. Neuropsychological measures of executive functions and episodic memory can identify the cognitive deficits associated with this pathology. Part of the utility of neuropsychological assessment lies in the fact that structural changes in late life depression are not uniform; some individuals will have higher lesion burden in prefrontal regions, while others will have greater subcortical pathology (9). Some individuals will experience considerable hippocampal atrophy, while others will not. Moreover structural abnormalities may be functionally benign in some individuals, but functionally disabling in others. Neuropsychological performance may provide a more sensitive indicator of neurobiological dysfunction and potential barriers to treatment response, which may also be complementary to structural findings.

Executive functions have received particular attention as a correlate of treatment response because worse performances in this domain have been associated with lower rates of remission in LLD. One study found response inhibition to predict 8-week treatment (22–23) and 4-month treatment response (24), while deficits in initiation and perseveration were associated with reduced 8- and 12-week treatment response (5,25). Executive functions purportedly rely on multiple pathways involving different regions of the prefrontal cortex and subcortical structures (25). As with mood, executive functions are similarly disrupted by cerebrovascular pathology, particularly in deep-white matter regions (26–27). Lesions in subcortical white matter have been associated specifically with slowed processing speed, which may be a mediator of cognitive deficits in the context of depression (28); however, research on treatment response has focused on only a few specific executive functions (e.g., response inhibition and initiation/perseveration), and consideration of additional variables (e.g., processing speed) may provide a more complete picture of this relationship.

The hippocampus also appears to play an important role in LLD, though how this association impacts treatment response remains unclear. Research has demonstrated that depressed older adults have reduced hippocampal volumes and corresponding deficits on tests of episodic memory (29). Reductions in hippocampal volume are also associated with longer duration of
untreated depression and higher recurrence (30–31), revealing variations in illness course that are associated with the hippocampus and may be important to predicting treatment response. In addition, smaller hippocampal volume at baseline has been associated with poorer response to treatment after one year in patients with LLD (32). Based on these findings, one might expect that deficits in episodic memory mediated by the hippocampus would be associated with reduced treatment response, though a clear link has yet to be demonstrated.

Thus, a better understanding of the relationship of executive and memory functions to treatment response in LLD may improve our ability to identify those individuals who are less likely to benefit from traditional pharmacological intervention. The goal of the current study was to further explore neuropsychological markers of treatment response in LLD. We included measures of attention, processing speed, and mental flexibility that are associated with frontostriatal function, and measures of episodic memory that are associated with hippocampal function. In this effort, we examined whether or not performance on neuropsychological measures would differentiate high and low treatment responders. It was hypothesized that patients with baseline deficits in memory and executive processes would show weaker response to treatment for depression than those with intact cognitive function. We predicted that these baseline cognitive deficits would be associated with poor antidepressant treatment response after controlling for the effects of current mood symptoms.

**Methods**

### Participants

Participants were enrolled in the NIMH-sponsored Mental Health Clinical Research Center (MHCRC). All participants were aged 60 or older and met DSM-IV criteria for Major Depressive Disorder upon enrollment. Diagnosis was determined by the treating psychiatrist using multiple clinical measures and following an established study protocol (33). Participants were excluded if they met criteria for a primary diagnosis of another major psychiatric illness, endorsed a history of alcohol or substance use problems, were diagnosed with a primary neurological condition, or were coping with other factors/conditions that may affect neuropsychological performance (e.g., barbiturate or hypnotic use, major medical illness with cognitive sequelae, significant sensory or motor limitations). Participants were excluded if they met clinical criteria for dementia at enrollment, or if they maintained a score below 25/30 on the Mini Mental State Examination (MMSE) across baseline and the initial 8 weeks of treatment.

### Materials

**Depression Assessment**—The Montgomery-Asberg Depression Rating Scale (MADRS) is a 10-item scale of depression severity that is based on patient report and clinical observation. This measure was designed to be sensitive to symptom change in clinical trials (34). Scores on the MADRS range from 0–60, with higher scores indicating greater severity of depressive symptomatology. Baseline and 1-year MADRS scores were obtained from the treating psychiatrist, who was blind to neuropsychological performance for each patient.

**Neuropsychological Functioning**—For the purposes of the current study, neuropsychological functioning was defined as performance on a brief battery of tests assessing memory and specific executive processes. This does not represent a comprehensive neuropsychological assessment, as several cognitive domains (e.g., visual-spatial and language functioning) were omitted. Memory and executive functioning were the focus of this study because previous research has identified these two domains as particularly affected in LLD. Additionally, the domain of executive functioning is somewhat broadly defined, and considerable variability exists in how researchers and clinicians operationalize this construct (35). We chose to include...
attention and processing speed in our definition of executive functions considering their purported reliance on prefrontal and subcortical structures, though we recognize that they may represent more basic executive skills than abilities associated with mental flexibility, response inhibition, and problem-solving.

Three measures comprised a brief assessment of executive functions for this study: Trail Making Test Part A (TMT-A), Trail Making Test Part B (TMT-B), and the Symbol Digit Modalities Test (SDMT). TMT-A assesses attention and visuomotor processing speed by requiring patients to quickly connect numbered circles in sequence. TMT-B assesses attention and visuomotor processing speed, but also demands mental flexibility by requiring patients to connect circles in alternating numerical and alphabetical sequences (36). Completion time in seconds was the outcome score for both TMT-A and TMT-B performances. The SDMT assesses attention, processing speed, and incidental memory (37). On the SDMT, patients use a key showing number and symbol pairs to write a series of numbers matching their corresponding symbols. The outcome score for the SDMT was total correct responses within 90 seconds.

Measures of delayed prose recall, prose percent retention, and delayed word-list recall comprised the memory assessment variables for this study. Specifically, the Logical Memory (LM) subtest of the Wechsler Memory Scale-Revised (WMS-R; 38) was administered to patients, which involved two orally presented narratives, an immediate recall trial, and a 30-minute delayed recall trial; the LM delayed recall score served as the variable of interest from these trials. LM percent retention was calculated as the number of correct details on the delayed recall trial divided by the number of correct details from the immediate recall trial. The final memory variable consisted of the number of correctly recalled words following a delay on the Consortium to Establish a Registry for Alzheimer’s Disease word-list (CERAD; 39). Ten words were orally presented to participants over three learning trials. Subjects were asked to recall as many words as possible following a 10-minute delay.

All participants completed this brief neuropsychological assessment battery at study entry and 1-year follow-up. The test battery was administered by a trained psychometrician who was blind to depression status.

Treatment Protocol

Participants received pharmacological treatment for depression using the Somatic Treatment Algorithm for Geriatric Depression protocol (STAGED, 33). The STAGED approach uses a standardized algorithm designed to simulate clinical treatment practices, factoring treatment history and symptom severity into dosing decisions. Patients with no history of treatment for depression were initially prescribed a selective serotonin reuptake inhibitor (SSRI). Those patients who continued to report moderate-to-severe symptoms of depression and/or no response to treatment after an SSRI trial of 8–12 weeks received either alternative treatment with venlafaxine, or augmentation of their original medication with buproprion. Patients who continued to show poor response to these treatment changes received additional alternatives, including tricyclic antidepressants or treatment augmentation with lithium. A detailed presentation of this medication algorithm appears in a previous publication (33).

Statistical analysis

Standard independent samples t-tests were used to compare groups along basic demographic variables. Analysis of Covariance (ANCOVA) was employed to compare patients responding to treatment for depression to those who did not respond. Baseline depression was treated as a covariate throughout analyses to control for variance associated with an acute mood disturbance. Similarly, depression scores at 1-year follow-up were also covaried when
appropriate cognitive functioning at follow-up was included in analyses. LM delayed recall, LM percent retention, delayed word-list recall, TMT-A, TMT-B, and SDMT scores served as dependent variables for separate ANCOVA models.

Results

The baseline sample (N = 177) included 110 Females (62%) and 67 Males (38%), with a mean age of 69.1 years (SD = 6.9) and mean education of 14 years (SD = 3.3). The mean baseline MADRS score for the entire sample was 21.8 (SD = 8.2).

One-hundred three patients were available for a follow-up assessment after one year. Those who did and did not return for follow-up did not differ in age (t [175] = 1.57, p = .106), education (t [175] = 1.29, p = .199), or baseline depression severity (t [171] = .229, p = .819). Patients were stratified by depression severity, defined as mild-to-moderate depression (MADRS < 26, n = 122) and moderate-to-severe depression (MADRS ≥ 26, n = 52). Patients in the moderate-to-severe group at baseline showed more improvement than the mild-to-moderate group in depressive symptoms at one-year follow-up (F (1, 99) = 25.27, p = .001, η² = .203, see Table 1). To examine differences at one-year follow-up, the sample was divided by treatment response, with improvement in depressive symptoms (high response) defined as a MADRS rating change greater than the baseline standard deviation for the sample. Demographic variables by treatment response groups appear in Table 1.

High response and low response groups did not differ in age (F (1, 98) = 2.79, p = .10, η² = .028) or education (F (1, 98) = .348, p = .557, η² = .004). Chi-square analyses showed no differences between groups for current psychotropic medication use (χ² (2) = .730, p > .05). Medication classifications and frequencies by group appear in Table 2. A small percentage of patients were not taking antidepressants because they were undergoing electroconvulsive therapy (ECT) while participating in this study. Rates of ECT did not differ between the high response and low response groups (χ² (1) = .099, p > .05).

Patients in the high response group at Year 1 demonstrated better baseline performance on several of the neuropsychological measures than the low response group. When covarying for the effects of baseline depression, high response individuals demonstrated better baseline LM delayed recall (F (1, 97) = 15.124, p = .001, η² = .135), higher LM percent retention (F (1, 97) = 9.56, p = .001, η² = .090), better delayed word-list recall (F (1, 97) = 5.210, p = .025, η² = .051), and higher SDMT scores (F (1, 93) = 5.951, p = .017, η² = .060) than low response participants. In contrast, the two patient groups did not differ significantly on TMT-A or TMT-B performances, though differences on the latter task were in the expected direction (F (1, 95) = 3.026, p = .085, η² = .031). Table 3 presents baseline performances for these measures.

Performance advantages for the high response group were maintained at 1-year follow-up when controlling for baseline depression alone. These differences, however, were no longer significant after covarying depression severity at follow-up, with the exception of TMT-B. Although TMT-B change scores were not significantly different after covarying depression severity at follow-up (F (1, 92) = .145, p > .05), the mean differences did suggest improved (faster) times on average for the high response group (M = −3.44, SD = 51.83) and worse (slower) times on average for the low response group (M = 7.88, SD = 39.18). Thus, differences in TMT-B performances at follow-up may reflect mild, but greater practice effects experienced by the high treatment response group than by the low response group. Scores and statistical results for testing at one-year follow-up are presented in Table 4.
Discussion

The current study found that depressed older adults with poor baseline performances on tests of verbal memory and processing speed demonstrated less treatment response over one year than patients with better cognitive test results. Severity of depression at baseline did not appear to influence the relationship between baseline neurocognitive function and the magnitude of treatment response. Participants in the high response group on average had MADRS scores falling within the remitted range, while the low response group on average remained in the mildly depressed range. Interestingly, patients with more severe depression demonstrated greater improvement over one year than patients with less severe symptoms, though again, depression severity was not a factor in baseline neurocognitive performance.

The current study remains consistent with previous research suggesting that deficits in executive function are associated with lower remission rates among depressed older adults (5,22–23,40–43). The current results particularly highlight the role of processing speed (SDMT performance), which emerged as the strongest putative frontostriatal indicator of treatment response. Consistent with this finding, previous research has suggested that cognitive slowing may reflect a more subcortical profile of depression in some patients (44).

Of all cognitive variables, baseline LM delayed recall produced the largest effect sizes between treatment response groups. Multiple prior studies have documented verbal memory deficits in LLD, with patients often showing some memory improvement following treatment with antidepressants (45–51). The current findings, however, provide somewhat unique evidence indicating that pre-treatment verbal memory may be a useful predictor of response to treatment. This is consistent with structural imaging studies that have tied reduced hippocampal volume to symptom duration in chronic depression (19–20), though episodic memory deficits in depression can also occur as a result of subcortical vascular pathology (26). The current results do not localize neuropsychological deficits in processing speed or episodic memory to specific structural abnormalities in frontostriatal pathways or the hippocampus. Rather, these findings highlight that neuropsychological correlates of treatment response extend beyond previous findings of initiation and perseveration, and provide additional evidence that treatment response may be complicated by heterogeneous etiologies in LLD.

The current study builds upon existing literature with the inclusion of a 1-year follow-up interval relative to previous 8–16 week intervals, assessment of neuropsychological performance at both baseline and follow-up, and control for the effects of depression severity at both baseline and follow-up assessments. In clinical settings, brief neurocognitive testing may assist treatment providers with clinical judgments and prognosis regarding pharmacological intervention for depression in older adults. While a comprehensive neuropsychological battery provides important information about the extent and nature of cognitive dysfunction, this study suggests that even a brief battery (roughly 35 minutes) could inform treatment practice and outcome.

Some features of this study’s design and results may limit our ability to generalize these findings. For instance, this design did not include a waitlist control group or single agent design, so we do not have data concerning cognitive functioning and natural remission in LLD. While we do not address natural symptom remission, the naturalistic treatment approach in this study was designed to enhance external validity and application in standard treatment settings. As opposed to single agent designs, the STAGED approach does not strictly control the treatment agent and dosage; rather, the goal is to emphasize a specific agent across participants while maintaining flexibility that achieves an optimal treatment response for each individual. Each patient was treated with the same progressive algorithm until notable symptom response was evident. Therefore, patients with more severe depression at baseline were not necessarily
treated more aggressively than less depressed individuals. A naturalistic treatment design is potentially more ecologically valid relative to treatment in community settings, though these results are also complementary to research findings based on single agent designs.

The issue of clinical versus statistical significance is a concern for studies of this nature. Although performances were significantly different between the two groups, effect sizes (<.3) were small across all measures. This suggests that performance differences observed in clinical settings may be subtle in nature and not readily observable without sensitive testing; however, the relationship between group differences and response to treatment for depression indicated that even mild, relative memory and processing speed weaknesses may warrant consideration when treating LLD. We acknowledge that a comprehensive neurocognitive battery might provide a more complex and developed indication of how these groups differ. Nevertheless, these results indicate that brief assessments of memory and processing speed may have clinical utility when treating patients with LLD.

Acknowledgments

This research was supported by NIMH grants P50 MH60451, R01 MH54846, and K24 MH70027

References


Table 1

Demographic variables by depression improvement.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Sex</th>
<th>Education</th>
<th>Baseline MADRS</th>
<th>Year 1 MADRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>High response</td>
<td>67.6</td>
<td>M = 37</td>
<td>14.2</td>
<td>24.79</td>
<td>7.52</td>
</tr>
<tr>
<td>N = 66</td>
<td>(5.5)</td>
<td>F = 29</td>
<td>(2.7)</td>
<td>(7.16)</td>
<td>(6.64)</td>
</tr>
<tr>
<td>Low response</td>
<td>69.8</td>
<td>M = 17</td>
<td>14.7</td>
<td>16.02</td>
<td>15.35</td>
</tr>
<tr>
<td>N = 34</td>
<td>(7.5)</td>
<td>F = 17</td>
<td>(3.5)</td>
<td>(7.97)</td>
<td>(8.52)</td>
</tr>
</tbody>
</table>

*Values in parentheses represent standard deviations.*
### Table 2
Medications for High response and Low response participants

<table>
<thead>
<tr>
<th></th>
<th>Antidepressants Only*</th>
<th>Antidepressants Plus†</th>
<th>Non-antidepressants‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Response</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 66</td>
<td>23% (15)</td>
<td>68% (45)</td>
<td>9% (6)</td>
</tr>
<tr>
<td>Low Response</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 34</td>
<td>29% (10)</td>
<td>65% (22)</td>
<td>6% (2)</td>
</tr>
</tbody>
</table>

* Antidepressants = non-tricyclic antidepressants, SSRIs, SSNRIs, MAOIs
† Patients taking antidepressants plus non-antidepressant psychotropic medications
‡ Non-antidepressant psychotropic medications = antianxiety, benzodiazepines, tricyclic antidepressants, antipsychotics, and anticonvulsants, with no antidepressants
Table 3
Baseline neurocognitive performance of High Response versus Low Response.

<table>
<thead>
<tr>
<th>Test</th>
<th>High response N = 66 Mean (SD)</th>
<th>Low Response N = 34 Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMS-R LM Delayed Recall†</td>
<td>20.50 (8.56)</td>
<td>15.50 (8.74)</td>
</tr>
<tr>
<td>WMS-R LM Percent Retention *</td>
<td>.79 (.198)</td>
<td>.72 (.256)</td>
</tr>
<tr>
<td>CERAD Word-List Recall ‡</td>
<td>6.21 (2.02)</td>
<td>6.00 (2.06)</td>
</tr>
<tr>
<td>Trail Making Test Part A</td>
<td>47.17 (36.25)</td>
<td>44.44 (21.13)</td>
</tr>
<tr>
<td>Trail Making Test Part B</td>
<td>123.20 (69.88)</td>
<td>141.21 (75.32)</td>
</tr>
<tr>
<td>Symbol-Digit Modalities Test †</td>
<td>38.94 (11.03)</td>
<td>35.62 (10.45)</td>
</tr>
</tbody>
</table>

* Multiply value by 100 for percentage score

† WMS-R LMII: F (1, 97) = 15.124, p = .001, η² = .135; WMS-R LM Percent Retention: F (1, 97) = 9.56, p = .001, η² = .090; CERAD Word List Recall: F (1, 97) = 5.210, p = .025, η² = .051; Trails A: F (1, 98) = .783, p = .378, η² = .008; Trails B: F (1, 95) = 3.026, p = .085, η² = .031; SDMT: F (1, 93) = 5.951, p = .017, η² = .060.
Table 4
Year 1 neurocognitive test results covarying baseline and follow-up depression severity

<table>
<thead>
<tr>
<th>Test</th>
<th>High Response†</th>
<th>Low Response†</th>
<th>ANCOVA Results‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMS-R LM Delayed Recall</td>
<td>22.78 (9.16)</td>
<td>18.41 (9.95)</td>
<td>F (1, 94) = 3.7, p = .057, η² = .038</td>
</tr>
<tr>
<td>LM Percent Retention*</td>
<td>.831 (.20)</td>
<td>.746 (.26)</td>
<td>F (1, 94) = 2.09, p = .152, η² = .022</td>
</tr>
<tr>
<td>CERAD Word-List Recall</td>
<td>6.91 (1.99)</td>
<td>6.18 (2.17)</td>
<td>F (1, 96) = 3.51, p = .064, η² = .035</td>
</tr>
<tr>
<td>Trail Making Test Part A</td>
<td>43.77 (25.35)</td>
<td>46.26 (24.23)</td>
<td>F (1, 96) = 3.72, p = .057, η² = .037</td>
</tr>
<tr>
<td>Trail Making Test Part B</td>
<td>114.50 (59.71)</td>
<td>153.56 (86.21)</td>
<td>F (1, 92) = 4.21, p = .043, η² = .044</td>
</tr>
<tr>
<td>Symbol-Digit Modalities Test</td>
<td>38.67 (12.02)</td>
<td>36.06 (11.46)</td>
<td>F (1, 93) = 2.78, p = .099, η² = .029</td>
</tr>
</tbody>
</table>

* Multiply value by 100 for percentage score.
† Values in parentheses represent standard deviations.
‡ Values in parentheses represent degrees of freedom.