

Response Inhibition Predicts Poor Antidepressant Treatment Response in Very Old Depressed Patients

Joel R. Sneed, Ph.D., Steven P. Roose, M.D., John G. Keilp, Ph.D.,
K. Ranga Rama Krishnan, M.B., Ch.B.,
George S. Alexopoulos, M.D., Harold A. Sackeim, Ph.D.

Objective: *There have been mixed findings regarding the prognostic significance of age of onset, executive dysfunction, and hyperintensity burden on treatment outcome in late-life depression. Methods:* Growth curve models were fit to data from the only 8-week, double-blind, placebo controlled trial of citalopram (20–40 mg/day) in patients aged 75 years and older with unipolar depression. Baseline assessment included Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) (to determine age at onset), Stroop Color-Word Test (to assess the response inhibition component of execution dysfunction), and structural magnetic resonance imaging (to determine hyperintensity burden). **Results:** *In the citalopram condition, patients with response inhibition (most impaired quartile) scored higher at endpoint than those without response inhibition. There were no effects for age of onset or hyperintensity load on response in the citalopram condition. In the placebo condition, patients with early-onset depression had higher depression scores at endpoint than patients with late-onset depression. Conclusion:* Only response inhibition, a fundamental executive function, predicted poor treatment response to antidepressant medication. Although patients with response inhibition also showed deficits in reaction time, adjusting for reaction time in our final response inhibition model did not substantively change the findings. (Am J Geriatr Psychiatry 2007; 15:553–563)

Key Words: Vascular depression, late-onset depression, executive dysfunction, cerebrovascular disease, treatment outcome, geriatric depression

The vascular depression hypothesis initially emerged from the finding that patients with late-

onset major depression (LOD) had higher rates of hyperintensity (HI) burden on structural magnetic

Received March 13, 2006; revised October 17, 2006; accepted October 20, 2006. From Department of Psychiatry, Columbia University and the New York State Psychiatric Institute, New York, NY (JRS, SPR, JGK, HAS); Department of Psychiatry, Duke University, Durham, NC (KRRK); and Department of Psychiatry, Weill Medical College of Cornell University, Westchester, NY (GSA). Send correspondence and reprint requests to Joel R. Sneed, Ph.D., Department of Psychiatry, Columbia University, Department of Biological Psychiatry, NY State Psychiatric Institute, Unit 98, 1051 Riverside Dr., New York, NY 10032. e-mail: js2627@columbia.edu

© 2007 American Association for Geriatric Psychiatry

resonance imaging (MRI) compared to patients with early onset depression.¹⁻³ It was further observed that patients with LOD and HI burden showed deficits on tests of executive functions,³⁻⁵ a broad, multifaceted class of functions consisting of mental planning and organization, cognitive flexibility, set development and shifting, and error monitoring.⁶ Based on these observations, it was hypothesized that structural brain damage, resulting from ischemic cerebrovascular disease, creates a vulnerability to depression in late life that is characterized, but not limited to, deficits in executive functioning.

The vascular depression hypothesis has had considerable heuristic impact and generated considerable research examining the prognostic significance of its key features with respect to treatment outcome. Several studies have concluded that overall HI burden in particular structures is associated with poorer outcome with antidepressant treatment,^{1,7,8} whereas other studies have not supported such associations.⁹ Various components of the executive functions (e.g., initiation and perseveration of responses and freedom from interference) have been found to be more impaired in patients with inadequate response to antidepressant treatment¹⁰ and in patients with increased rates of relapse/recurrence during maintenance treatment,¹¹ but negative findings have also been reported.¹² There have been mixed results with regard to LOD as well. For example, LOD has been associated with slow recovery in some studies¹³ but other studies have failed to document this association.^{14,15}

The inconsistency in findings may be related to the major methodological limitations that have characterized much of this research. Treatment in most studies was unconstrained, administered in an open fashion, without standardization, or even restricted to a single medication or class of medication. Thus, it is difficult to characterize exactly what was being predicted, such as outcome to antidepressant treatment in general or to a specific mixture of possible antidepressant treatments. Furthermore, the three major indices of the vascular depression construct have rarely been assessed in the same patient sample and contrasted in predictive power with respect to clinical outcome. Consequently, it is unknown whether the predictive power of any index is unique or redundant with that of the other indices.

A large, double-blind, randomized clinical trial was conducted in patients with major depression 75 years of age and older contrasting the safety and efficacy of citalopram and placebo.¹⁶ This study provided the context for the present study in which we test the prognostic significance of LOD, response inhibition (a fundamental component of executive function), and HI burden. According to Alexopoulos,¹⁷ the poor response to antidepressant treatment observed in patients with deficits in executive functions is that ischemic lesions (or age-related brain changes) disrupt the frontal system (frontal basal ganglia-thalamocortical circuits), which in turn compromises the neural networks necessary for antidepressant response. For example, the prefrontal cortex is the major neocortical target for information processed in the limbic system.¹⁸ Thus, we hypothesized that patients with LOD, response inhibition (RI), or HI burden would respond less well to treatment than patients without LOD, RI, or HI burden in the citalopram condition. Conversely, because there is no developed theory about the mechanism of action of the placebo effect, we did not have any reason to predict differences between patients with and without LOD, RI, or HI burden in this sample.

In addition to testing these primary hypotheses, we conducted secondary analyses that focused on whether defining groups based on combinations of LOD, RI, and HI burden predicted treatment outcome in either condition, addressing the issue of unique versus redundant predictive value among these core features of the vascular depression hypothesis. Finally, because the criteria for defining LOD, RI, and HI burden were based on precedent set in the literature and not empirically derived, we conducted planned tertiary analyses exploring whether the findings from the primary analyses differed by using alternative cut points for LOD, RI, and HI burden.

METHODS

The procedures used in the multisite, randomized, placebo-controlled trial (RCT) have been previously described.¹⁶ Briefly, 174 community-dwelling men and women 75 years or older meeting *DSM-IV* cri-

teria (based on the Structured Clinical Interview for DSM-III-R [SCID]) for nonpsychotic unipolar depression (single or recurrent) with a baseline 24-item Hamilton Rating Scale for Depression (HRSD) score ≥ 20 participated in this eight-week RCT. All patients began the trial with a one-week single-blind placebo lead-in with the baseline visit conducted at the end of the lead-in. Patients were randomized to citalopram 20 mg/day or matched placebo only if they continued to meet inclusion and exclusion criteria at the end of the placebo lead-in. At the end of week 4, patients with a HRSD score >10 had the dose increased to two pills per day: 40 mg of citalopram or two placebo pills.

Age at onset was determined based on information obtained in the baseline SCID interview, medical records, and other sources of historical information. Currently, there is no consensus on whether age of onset is best treated as a continuous or categorical variable and, if categorical, on the threshold value that distinguishes early-onset depression (EOD) and LOD. With respect to categorical approaches, some investigators have used an age threshold of 50 years or greater to define LOD,^{1,5,10,19,20} whereas others have used an age threshold of 60 years or greater.^{3,4,21} Especially in light of the age range of the present sample, patients were considered to have LOD if they experienced their first episode of major depression after 60 years of age. Patients with onset earlier than 60 years of age were considered as manifesting EOD.

Response inhibition, a fundamental executive function,²² was measured using the Stroop Color-Word Test.²³ The Stroop presents stimuli individually, allowing reaction time to be recorded for each trial. By key press, subjects identify on each trial the name or color of a stimulus. There are three blocks of trials, each with a 50 msec inter-stimulus interval. In the first block (word-only condition), printed color names (red, blue, green) are presented in black. In the second block (color-only condition), a string of the letter X is presented in different colors (red, blue, green). In the third block, color names are again presented, but this time printed in incongruous colors. Subjects are given feedback on each trial (beep for correct; buzz for error), and this task produces consistent interference effects. Patients were classified as having RI if their response inhibition scores

were in the highest quartile of the distribution, consistent with previous definitions.²⁴

HI burden was determined on the basis of T1-weighted, proton-density, and T2-weighted images evaluated for the presence of deep-white-matter and subcortical gray-matter lesions. The severity of lesions was graded by a neuroradiologist using the Fazekas modified Coffey Rating Scale for signal hyperintensities.² Deep white matter hyperintensities (DWMH) were defined as abnormalities in the frontal, parietal, temporal, or occipital lobes, and scored as 0 (absent), 1 (punctate foci), 2 (beginning confluence of foci), and 3 (large confluent areas). Subcortical gray matter lesions were defined as abnormalities in the caudate nucleus, putamen, globus pallidus, thalamus, and internal capsule, and were scored as 0 (absent), 1 (punctate), 2 (multipunctate), and 3 (diffuse). The readers were blinded to treatment condition. In this study, patients were classified as HI burden on the basis of Krishnan's criteria of a grade of 2 or more on DWMH or a score of 3 in the subcortical gray rating.²

Data Analysis

The mixed models procedure in SPSS version 14 (SPSS, Inc., Chicago, IL) was used to fit a series of growth curve models in the intent-to-treat citalopram (N = 84) and placebo (N = 90) samples separately. We fit these models separately because it was not of substantive interest to compare response across treatment conditions, such as comparing differences at endpoint across treatment conditions in patients with LOD, RI, or HI burden.

Mixed effects or growth curve models have a number of advantages over more traditional analyses of repeated measures data that allow for fuller exploitation of the available data. They accommodate time-varying repeated measures, permit inclusion of individuals not assessed at all time points, are tolerant of unequal intervals between data points, and time can be treated flexibly so that the intercept can correspond to any time point in the study.²⁵ Growth curve models, which make use of maximum likelihood estimation, are also an improvement over endpoint analyses based on last observation carried forward (LOCF) data. LOCF makes the unlikely assumption that data are missing completely at random and patient response is constant from the last observation to

the end of the trial, assumptions that are likely to lead to biased results.²⁶

Growth curves provide estimates of both fixed effects (i.e., average effects of predictors on depression across participants) and random effects (i.e., variation in individuals' means and slopes). All models reported here included random effects for the intercept and linear and quadratic changes in depression with time. In addition, we treated the variance-covariance matrix as unstructured allowing for correlation among these random components. However, as the current study focuses on mean level change in depression over time, and whether this change differs depending on LOD, RI, or HI burden status, only the fixed effects are reported.

To determine the prognostic significance of LOD, RI, and HI burden in predicting treatment outcome in either condition, we began with the full interaction model. This model included the intercept, the main effects of time and time squared, a patient group term (LOD, RI, or HI burden), and its interaction with time and time squared. We used a "top-down" approach in which the highest order term was removed on each successive run if nonsignificant ($p < 0.05$) until the baseline model was reached. We considered the baseline model to include time, time squared, and the main effect of the patient group term.

All models adjusted (covaried) for baseline depression severity, site of study, which was found to be an important predictor in previous research,¹⁶ age, and gender (all centered at their respective means). Dummy variable coding was used for LOD (1 = LOD, 0 = EOD), RI (1 = RI, 0 = no RI), and HI burden (1 = burden, 0 = no burden). In addition, time was

centered at week 8 because we were interested in depression severity scores at endpoint.²⁵ Thus, the intercept reflected the average level of depression at week 8 for patients who were not classified as having LOD, RI, or HI burden (depending on the model).

RESULTS

Sample Characteristics

As previously reported (16), the average study participant was 79 years old and completed 1–2 years of college education. The sample consisted of 58% women with more women being in the placebo condition (62.2%) than in the citalopram condition (53.6%). However, this difference was not statistically significant ($\chi^2 = 1.34, p = 0.25$). Depression severity for the sample was moderate (average 24-item HRSD = 24). The average age of onset of depression was approximately 68; consequently, the majority of patients were classified as LOD (75.9%). Similarly, most were classified as having HI burden (72.6%), but by definition, only 22.4% of patients met criteria for RI. There was no association between patients with LOD and HI burden ($\chi^2 = 0.006, p = 0.94$), LOD and RI ($\chi^2 = 1.63, p = 0.20$), or HI burden and RI ($\chi^2 = 1.26, p = 0.26$).

Primary Hypotheses

The first set of hypotheses tested was: 1) patients classified with LOD would respond less well to antidepressant treatment than patients classified with EOD, and 2) there would be no difference in outcome

TABLE 1. Final Growth Curve Model Estimates for Late-Onset Depression (LOD) in the Citalopram and Placebo Conditions

Parameter	Citalopram						Placebo					
	B	SE(B)	t	df	p	95% CI	B	SE(B)	t	df	p	95% CI
Intercept	13.51	1.22	11.12	111	0.001	11.11, 15.92	19.52	1.65	11.81	80.11	0.001	16.23, 22.81
Time	-0.33	0.40	-0.81	67.13	0.42	-1.13, 0.48	0.24	0.38	0.63	131.52	0.53	-0.52, 1.00
Time ²	0.15	0.059	2.52	72.54	0.014	0.032, 0.27	0.12	0.049	2.50	84.43	0.014	0.025, 0.22
LOD	-0.29	1.14	-0.25	80.21	0.80	-2.55, 1.97	-6.13	1.88	-3.26	79.94	0.002	-9.87, -2.38
LOD × time							-0.66	0.31	-2.14	81.42	0.036	-1.27, -0.045
-2 log likelihood	2,946.86						3,296.83					
Parameters	15						16					

Notes: Dependent variables are the Hamilton Rating Scale for Depression (HRSD); RI (0 = early-onset depression, 1 = late-onset depression); time is centered at week 8. All models covary for baseline HRSD, site, age, and gender centered at their means (not tabled).

B: unstandardized regression weight; SE: standard error; CI: confidence interval.

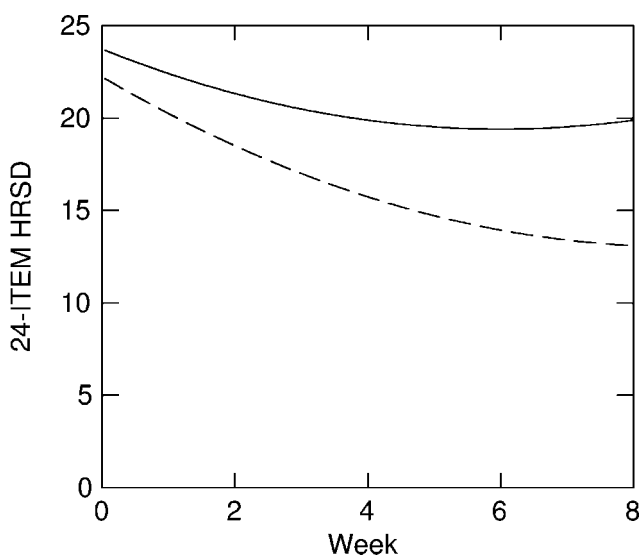
among placebo patients with and without LOD. As can be seen from Table 1, LOD did not predict outcome in the citalopram condition. In the placebo sample, our final model (Table 1) revealed a significant $\text{LOD} \times \text{Time}$ interaction indicating that depression scores for EOD patients did not decrease as sharply ($B = 0.24$) as depression scores for LOD patients ($B = 0.24 - 0.66 = -0.42$). This resulted in a significant difference at endpoint between patients with EOD ($M = 19.52$) and patients with LOD ($M = 19.52 - 6.13 = 13.39$), which corresponded to a large standardized effect (Cohen's $d = 0.83$; Figure 1).

The second set of hypotheses tested was: 1) patients classified with RI would respond less well to antidepressant treatment than patients classified as not having RI, and 2) there would be no difference between patients classified as having RI compared to patients classified as not having RI in the placebo condition. Consistent with expectations, our final model (Table 2) revealed a significant $\text{RI} \times \text{Time}$ interaction in the citalopram sample indicating that depression scores for patients classified as having RI did not decrease as sharply ($B = -0.50 + 0.68 = 0.18$) as depression scores for patients classified as not

having RI ($B = -0.50$). This difference in linear slopes resulted in a significant difference at endpoint between the two groups of patients. Specifically, patients with RI scored 5.95 points higher at endpoint than patients without RI ($M = 12.39$). This endpoint difference corresponded to a large standardized effect (Cohen's $d = 0.86$; Figure 2). As can be seen from Table 2, RI did not predict outcome in the placebo condition.

The third set of hypotheses to be tested was: 1) patients with HI burden will respond less well to antidepressant treatment than patients without HI burden, and 2) there would be no difference between patients with and without HI burden among patients in the placebo condition. Contrary to expectations, HI burden did not predict outcome in the citalopram (Table 3). There was, however, a significant $\text{HI burden} \times \text{Time}$ interaction in the placebo sample indicating that depression scores for patients with HI burden did not decrease as sharply ($B = -0.76 - 0.65 = -0.11$) as depression scores for patients without HI burden ($B = -0.76$). This difference in the linear slopes, however, did not result in a significant difference at endpoint.

FIGURE 1. Growth Curves of Average Hamilton Rating Scale for Depression Scores Over Time for Late-Onset Depressed Patients (dashed line) and Early-Onset Depressed Patients (solid line) in the Placebo Condition



Secondary Analyses

To address the issue of unique versus redundant predictive value among the core features of the vascular depression hypothesis, we conducted secondary analyses in which we used different combinations of LOD, RI, and HI burden to classify patients (e.g., LOD and RI, LOD and HI burden, RI and HI burden). The only classification that predicted outcome included patients with LOD and RI in the citalopram sample. However, the final model for this combined patient group was not substantively different from the final model for the RI only model in the citalopram sample. In contrast to patients without LOD and RI ($B = 12.70$), patients with LOD and RI scored 5.64 points higher at endpoint, $t(67.085) = 2.55$, $p < 0.013$, 95% CI = 1.22, 10.07 (Table 2). Because LOD alone does not predict treatment outcome in the citalopram condition, and when combined with RI does not yield a model that is substantively different from the one obtained using RI only, LOD does not appear to explain unique variance in predicting treatment outcome.

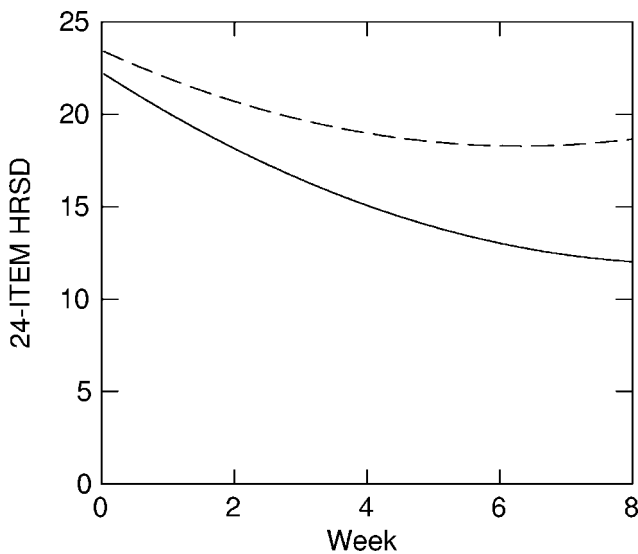
TABLE 2. Final Growth Curve Model Estimates for Response Inhibition (RI) in the Citalopram and Placebo Conditions

Parameter	Citalopram						Placebo					
	B	SE(B)	t	df	p	95% CI	B	SE(B)	t	df	p	95% CI
Intercept	12.39	1.00	12.45	66.37	0.001	10.40, 14.38	15.020	0.86	17.45	82.86	0.001	13.31, 16.73
Time	-0.50	0.45	-1.10	66.20	0.28	-1.40, 0.40	-0.28	0.31	-0.90	78.38	0.37	-0.90, 0.34
Time ²	0.13	0.065	2.057	66.86	0.044	0.0039, 0.26	0.11	0.050	2.27	79.38	0.026	0.013, 0.21
RI	5.95	1.90	3.14	69.74	0.002	2.17, 9.73	-1.94	1.15	-1.69	82.26	0.096	-4.23, 0.35
RI × time	0.68	0.33	2.089	66.59	0.041	0.03, 1.33						
-2 log likelihood	2,694.72						3,148.83					
Parameters	16						15					

Notes: Dependent variables are the Hamilton Rating Scale for Depression (HRSD); LOD (0 = RI, 1 = no RI); time is centered at week 8. All models covary for baseline HRSD, site, age, and gender centered at their means (not tabled).

B: unstandardized regression weight; SE: standard error; CI: confidence interval.

FIGURE 2. Growth Curves of Average Hamilton Rating Scale for Depression (HRSD) Scores Over Time for Patients Classified as Having Response Inhibition (dashed line) and Not Having Response Inhibition (solid line) in the Citalopram Condition



Tertiary Analyses

Additional analyses determined whether alternative cut points to define the subgroups on LOD, RI, and HI burden would result in different findings from the primary analyses. LOD (redefined as age at onset ≥50) did not predict outcome in either sample. Redefining RI as corresponding to Stroop scores above or below the mean resulted in a significant

baseline model in which patients classified as having RI in the placebo condition scored on average 2.34 points lower than patients classified as not having RI in the placebo condition over the course of trial ($t(81.58) = -2.28, p = 0.025, 95\% \text{ CI} = -4.38 \text{ to } -0.30$). The above finding simply reflects an overall difference between RI and non-RI patients that existed at the beginning of the trial in the placebo condition. Redefining HI burden as corresponding to deep white matter or subcortical gray matter lesion scores of 3 did not predict treatment outcome in the citalopram condition and yielded a model that was substantively equivalent to the findings observed in the primary analyses for the placebo sample (Table 3).

Post-hoc Analyses

The Stroop Color-Word Test is a complex neuropsychological task that focuses on the slowing of reaction time in the color-word conflict condition relative the color-only condition. As opposed to a specific deficit in response inhibition, it is possible that a generalized slowing of reaction time or a slowing that occurs with greater cognitive demand accounted for the relationship between baseline Stroop performance and antidepressant response. To determine the nature of the baseline Stroop impairment, we contrasted the patients classified as having RI with the remaining patients in median reaction time (log transformed) on a number of tasks in the baseline battery. All post-hoc tests were evaluated for significance at the more conservative 1% level; point estimates with associated significance levels at or below 0.05 were considered marginally significant.

Within the Stroop, the groups did not differ in

TABLE 3. Final Growth Curve Model Estimates for Hyperintensity (HI) Burden in Citalopram and Placebo Conditions

Parameter	Citalopram						Placebo					
	B	SE(B)	t	df	p	95% CI	B	SE(B)	t	df	p	95% CI
Intercept	13.37	1.37	9.75	81.096	0.001	10.64, 16.097	12.42	1.94	6.40	55.069	0.001	8.53, 16.31
Time	-0.37	0.40	-0.93	47.99	0.36	-1.17, 0.43	-0.76	0.43	-1.74	83.79	0.085	-1.62, 0.11
Time ²	0.14	0.059	2.41	51.026	0.02	0.023, 0.26	0.13	0.064	1.97	55.45	0.052	-0.001, 0.25
HI	0.099	1.32	0.075	55.92	0.94	-2.55, 2.74	2.92	2.27	1.29	53.74	0.20	-1.63, 7.47
HI × time							0.65	0.31	2.087	53.38	0.042	0.025, 1.27
-2 log likelihood	2,179.90						2,176.46					
Parameters	15						16					

Notes: Dependent variables are the Hamilton Rating Scale for Depression (HRSD); HI (0 = no HI, 1 = HI); time is centered at week 8. All models covary for baseline HRSD, site, age, and gender centered at their means (not tabled).

B: unstandardized regression weight; SE: standard error; CI: confidence interval.

reaction time in the color-only condition ($t(156) = 0.88, p = 0.38$) but there was a marginally significant difference in reaction time in the word-only condition ($t(155) = 1.94, p = 0.054$) and a marked difference in reaction time in the color-word conflict condition ($t(156) = 9.71, p < 0.0001$). This indicates that RI is not attributable to differences in gross reaction time but to differences in reaction time in a conflict condition, which is reflected in the response inhibition or interference score. On more complex cognitive tasks, patients with RI had slower reaction times for correct responses on the judgment of line orientation (JOLO) test²⁷ ($t(145) = 2.65, p = 0.009$) and there was a trend for these patients to have slower reaction times to correct responses on the choice reaction time (CRT) test²⁸ ($t(153) = 2.28, p = 0.024$). Thus, those with Stroop impairment (RI) show deficits in speed of processing that involves more cognitive load, such as spatial components or complex visual arrays. However, including reaction time to correct responses on either the CRT or JOLO in the final RI model (Table 2) did not eliminate the effect of Stroop impairment on treatment outcome observed in the citalopram sample. Thus, the effect of response inhibition (Stroop impairment) on treatment outcome in the citalopram sample is independent of reaction time on complex tasks.

The issue of specificity was also tested by contrasting whether the groups defined by the Stroop interference effect differed in performance accuracy in the other neuropsychological measures included in the baseline battery. The groups were compared in global cognitive status, assessed by the 30-item Mini-Mental State Exam,²⁹ psychomotor performance using the digit-symbol test,³⁰ visual spatial organiza-

tion using accuracy scores on the JOLO, and verbal learning and memory using total words recalled in the learning and delayed phases of the Buschke Selective Reminding Test (SRT).³¹ No differences between those with and without impaired Stroop performance were observed.

An additional possibility is that patients with Stroop impairment have poorer response to antidepressant medication because they have greater medical burden than patients without Stroop impairment. However, there were no differences between patients with and without Stroop impairment in physician-rated chronic medical illness burden ($t(156) = 0.33, p = 0.74, 95\% \text{ CI} = -1.16 \text{ to } 1.62$) as assessed by the Cumulative Illness Rating Scale—Geriatrics.³²

DISCUSSION

This study examined whether patients with LOD, RI, and/or HI burden responded less well to antidepressant medication (citalopram) than patients without LOD, RI, and/or HI burden in an eight-week, randomized, placebo-controlled trial in patients at least 75 years of age. Only RI (defined as the highest quartile on the Stroop) identified a group of patients with poor clinical outcome in the citalopram condition. Neither LOD nor HI burden influenced treatment outcome in the citalopram condition. The only significant effect in the placebo condition was that patients with EOD had a significantly higher HRSD score than their LOD counterparts at the end of treatment.

This study is unique in several ways. The context for this study was the first double masked, randomized trial comparing a selective serotonin reuptake inhibitor to placebo in “old-old” depressed patients (75 years of age and older). Moreover, this was the only study in this age range in which age-at-onset, a component of the executive functions, and MRI were assessed in the same sample allowing for the simultaneous assessment of three potentially interrelated indices. Thus, we were able to determine the differential impact of age-of-onset, RI, and MRI lesion load on treatment response.

Patients classified as having RI had higher endpoint HRSD scores than patients classified as not having RI. These findings converge with an emerging literature documenting the impact of response inhibition, a fundamental component of executive function, on antidepressant treatment response in late-life depression. For example, Stroop scores at the most impaired quartile predicted lower remission rates in elders with major depression treated with citalopram at a target dose of 40 mg daily.²⁴ In another study, those who remained symptomatic showed greater deficits on the initiation/perseveration subtest of the Mattis Dementia Rating Scale as compared to patients who achieved remission.³³ Most recently, reaction time in the conflict condition of the Attention Network Test (a similar task to the Stroop) predicted prolonged time to remission of late-life depression.³⁴

The Stroop impairment observed in this study was not attributable to differences in simple reaction time but to differences in reaction time in a conflict condition, which is reflected in the response inhibition score on which the classification of patients was based. There were, however, differences in reaction time on more complex tasks (tests involving spatial components or complex visual arrays). These differences in reaction time are consistent with the critical role attributed to processing speed in mediating cognition.³⁵ According to this view, processing speed affects performance on complex (i.e., executive) tasks because: 1) they are dependent on simpler operations, and therefore, must be completed very rapidly; and 2) deficits in processing speed limits the amount of information simultaneously available. Subsequent research in late-life depressed versus normal control patients has shown that the cognitive deficits observed across a broad array of neuropsychological

tests in late-life depressed patients can be almost entirely explained by differences in processing speed.³⁶ However, including reaction time to correct responses on more complex tasks (e.g., JOLO and CRT) as covariates in the final RI model did not eliminate the differences in linear slopes and endpoint means observed between those classified as having and not having Stroop impairment. This indicates that the poor response to citalopram observed in patients with RI is over and above the deficits in reaction time observed on more complex tasks.

We did not find a relationship between MRI hyperintensity load and treatment response in this sample. The findings from this study may differ from other investigations^{2,37} for a number of reasons. The sample had a higher rate of hyperintensity burden (approximately 75%) and a higher mean age for the total sample (79.6 years). It is possible that the higher rate of hyperintensity load observed in this study as a function of age obscures the potential association between hyperintensity burden and treatment response. As a result, the utility of MRI in the diagnosis of patients with vascular depression may be restricted to early late-life (e.g., mean age 60) when lesions are more specific to the vascular depression syndrome.

It could also be argued that the lack of association between hyperintensity load and response to citalopram is due to the low response rate of this very old sample. However, there was an approximately 40% response rate in this study¹⁶ and there was no association between patients with and without high hyperintensity burden and responder status in either treatment condition. Another reason for the lack of association between hyperintensities and response to citalopram may be that hyperintensities, as in other studies,² were not coded with respect to hemisphere or lobe potentially obscuring the relationship between lesion load in the frontal system and antidepressant response.

Late-onset depression also did not influence treatment outcome in either condition. Methodologically, onset of depression is difficult to identify, especially when early episodes are of mild severity and difficult to recall.^{37,38} Conceptually, neurological changes may contribute to a late-life episodes regardless of other depressive episodes in early life. Moreover, early onset depression may be a risk factor for late-

life depression by contributing to brain abnormalities predisposing to depression. One such mechanism, for example, involves stress-related hormones leading to reduction of neurotrophic factor secretion, and ultimately decreasing neurogenesis at the dentate nucleus of the hippocampus. These methodological and conceptual concerns difficulties have lead researchers^{37,39} to suggest that the dichotomous classification of depression based on age at onset may be inadequate.

Patients with early onset depression had significantly higher mean endpoint HRSD scores than their LOD counterparts in the placebo condition. One explanation for this effect is that older patients with early onset depression have a higher number of episodes of major depression than patients with late-onset depression, and recurrent depression is less likely to respond to treatment than single depressive episodes. These findings are consistent with data from the sertraline antidepressant heart attack randomized trial (SADHART) study in which patients in the placebo condition with recurrent depression had significantly less change in depression severity scores from baseline to endpoint than patients in a first episode.⁴⁰ In the present study, 94.6% of the patients classified as EOD had recurrent depressive episodes, whereas only 24.2% of patients classified as LOD had recurrent depression ($\chi^2 = 59.78$, $p < 0.001$).

This study has several limitations that are balanced by its unique methodological strengths. We only measured one component of the executive functions: response inhibition. As previously indicated, the executive functions consist of mental planning and organization, novel problem solving or cognitive flexibility, set development and shifting, and error monitoring.⁶ It is unclear whether the findings reported in this study generalize beyond response inhibition to other aspects of executive functioning. Similarly, the choice to examine the impact of response inhibition on treatment outcome was based on the vascular depression hypothesis and does not reflect the distinct possibility that other aspects of cognition predict poor treatment response in the very old depressed. The specificity of the relationship between response inhibition and depression, therefore, remains unclear and should be examined in future research. The hyperintensity rating method

may also have limitations. In particular, it may be limited because of the age range of the sample and the degree of abnormality that exists. Future studies should also use volumetric measurement of hyperintensities and code hyperintensities taking into account lobe and hemisphere location. Specificity with respect to lobe and hemisphere is especially important given the critical role of the prefrontal cortex in mediating the executive functions.¹⁸

This study examined the prognostic significance of the core features of the vascular depression hypothesis on treatment outcome. We found that only response inhibition, a fundamental executive function, predicted poor response in the citalopram condition, suggesting that future therapeutic efforts focus on approaches that aim to alleviate deficits in executive control such as problem solving therapy.⁴¹ The variables that effect treatment outcome are not only clinically important but they are also relevant to the diagnostic criteria that underlie proposed late-life constructs such as vascular depression or the depression executive dysfunction syndrome.^{17,42} As Sneed et al.⁴³ have argued, it is unclear whether these two proposed entities are simply variations of a single underlying disorder or whether they define distinct patient groups, and whether the core features of these two proposed constructs (e.g., late-onset depression, executive dysfunction, and MRI hyperintensities) are unique or redundant.

Although the present findings provide support for the depression executive dysfunction syndrome hypothesis,^{17,42} we caution researchers against generalizing from a single test like the Stroop to a domain as broad as the execution functions. By simultaneously assessing age of onset, executive dysfunction, and hyperintensity load in the same sample, the current study addresses critical issues in late-life depression treatment outcome research. It also raises a number of psychometric questions regarding the assessment of late-onset depression, executive dysfunction, and MRI hyperintensities that need to be systematically addressed in order to evaluate the validity of proposed diagnostic entities in late-life.

This research was supported by a grants from Forest Laboratories and the National Institute for Mental Health (T32 MH20004 to SPR and K23 MH075006 to JRS).

References

1. Hickie I, Scott E, Mitchell P, et al: Subcortical hyperintensities on magnetic resonance imaging: clinical correlates and prognostic significance in patients with severe depression. *Biol Psychiatry* 1995; 37:151-160
2. Krishnan KR, Hays JC, Blazer DG: MRI-defined vascular depression. *Am J Psychiatry* 1997; 154:497-501
3. Salloway S, Malloy P, Kohn R, et al: MRI and neuropsychological differences in early- and late-life-onset geriatric depression. *Neurology* 1996; 46:1567-1574
4. Alexopoulos GS, Meyers BS, Young RC, Kakuma T, Silbersweig D, Charlson M: Clinically defined vascular depression. *Am J Psychiatry* 1997; 154:562-565
5. Lesser I, Boone K, Mehlinger C, Wohl M, Miller B, Berman N: Cognition and white matter hyperintensities in older depressed patients. *Am J Psychiatry* 1996; 153:1280-1287
6. Podell K, Lovell MR: Neuropsychological assessment, in *Textbook of Geriatric Neuropsychiatry*. Edited by Coffey CE, Cummings JL. Washington DC, American Psychiatric Press, 2000
7. Simpson S, Baldwin RC, Jackson A, Burns AS: Is subcortical disease associated with a poor response to antidepressants? Neurological, neuropsychological, and neuroradiological findings in late-life depression. *Psychol Med* 1998; 28:1015-1026
8. Taylor WD, Steffens DC, MacFall JR, et al: White matter hyperintensity progression and late-life depression outcomes. *Arch Gen Psychiatry* 2003; 60:1090-1096
9. Salloway S, Correia S, Boyle P, et al: MRI subcortical hyperintensities in old and very old depressed outpatients: the important role of age in late-life depression. *J Neurol Sci* 2002; 203:227-233
10. Baldwin R, Jeffries S, Jackson A, et al: Treatment response in late-onset depression: relationship to neuropsychological, neuro-radiological and vascular risk factors. *Psychol Med* 2004; 34:125-136
11. Alexopoulos GS, Meyers BS, Young RC, et al: Executive dysfunction and long-term outcomes of geriatric depression. *Arch Gen Psychiatry* 2000; 57:285-290
12. Butters MA, Bhalla RK, Mulsant BH, et al: Executive functioning, illness course, and relapse/recurrence in continuation and maintenance treatment of late-life depression. *Am J Geriatric Psychiatry* 2004; 12:387-394
13. Alexopoulos GS, Meyers BS, Young RC, et al: Recovery in geriatric depression. *Arch Gen Psychiatry* 1996; 53:305-312
14. Driscoll HC, Basinski J, Mulsant BH, et al: Late-onset major depression: clinical and treatment-response variability. *Int J Geriatric Psychiatry* 2005; 20:661-667
15. Reynolds CF3rd, Dew MA, Frank E, et al: Effects of age at onset of first lifetime episode of recurrent major depression on treatment response and illness course in elderly patient. *Am J Psychiatry* 1998; 155:795-799
16. Roose SP, Sackeim HA, Krishnan KRR, et al. Group O-ODS: Antidepressant pharmacotherapy in the treatment of depression in the very old: a randomized, placebo-controlled trial. *Am J Psychiatry* 2004; 161:2050-2059
17. Alexopoulos GS: "The depression-executive dysfunction syndrome of late life": a specific target for D3 agonists? *Am J Geriatric Psychiatry* 2001; 9:22-29
18. Royall DR, Lauterbach EC, Cumings JL, et al: Executive control function: a review of its promise and challenges for clinical research. *J Neuropsychiatry Clin Neurosci* 2002; 14:377-405
19. Krishnan KRR, Hays JC, Tupler LA, George LK, Blazer DG: Clinical and phenomenological comparisons of late-onset and early-onset depression. *Am J Psychiatry* 1995; 152:785-788
20. Steffens DC, Krishnan KR: Structural neuroimaging and mood disorders: recent findings, implications for classification, and future directions. *Biol Psychiatry* 1998; 43:705-712
21. Van den Berg MD, Oldehinkel AJ, Bouhuys AL, Brilman EI, Beekman AT, Ormel J: Depression in later life: three etiologically different subgroups. *J Affective Disorders* 2001; 65: 19-26
22. Lezak MD: *Neuropsychological Assessment*. New York, Oxford University Press, 1995
23. Golden CJ: *The Stroop Color and Word Test: A Manual for Clinical and Experimental Use*. Wood Dale, Stoelting, 1978
24. Alexopoulos G, Kiosses DN, Murphy CF, Shanmugham B, Gunning-Dixon F: Executive dysfunction and the course of geriatric depression. *Biol Psychiatry* 2005; 58:204-210
25. Singer JD, Willett JB: *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York, Oxford University Press, 2003
26. Mallinckrodt CH, Kaiser CJ, Watkin JG, Molenberghs G, Carroll RJ: The effect of correlation structure on treatment contrasts estimated from incomplete clinical trial data with likelihood-based repeated measures compared with last observation carried forward. *Clin Trials* 2004; 1:477-489
27. Benton AL, Sivan AB, Hamsher KD, Varney N, Spreen O: *Contributions to Neuropsychological Assessment*. New York, Oxford University Press, 1994
28. Hindmarch I: Relevant psychometric tests for antidepressants and anxiolytics. *Int Clin Psychopharmacol* 1994; 9:27-33
29. Folstein MF, Folstein SE, McHugh PR: "Mini-Mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatric Res* 1975; 12:189-198
30. Wechsler D: *Wechsler Adult Intelligence Scale-III*. San Antonio, Psychological Corporation, 1997
31. Buschke H, Fuld PA: Evaluation of storage, retention, and retrieval in disordered memory and learning. *Neurology* 1974; 24:1019-1025
32. 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; 41:237-248
33. Kalayam B, Alexopoulos GS: Prefrontal dysfunction and treatment response in geriatric depression. *Arch Gen Psychiatry* 1999; 56:713-718
34. Murphy GF, Alexopoulos GS: Attention network dysfunction and treatment of response of geriatric depression. *J Clin Exp Neuropsychol* 2006; 28:96-100
35. Salthouse TA: The processing-speed theory of adult age differences in cognition. *Psychological Rev* 1996; 103:403-428
36. Butters MA, Whyte EM, Nebes RD, et al: The nature and determinants of neuropsychological functioning in late-life depression. *Arch Gen Psychiatry* 2004; 61:587-595
37. Krishnan KR, Taylor WD, McQuoid DR, et al: Clinical characteristics of magnetic resonance imaging-defined subcortical ischemic depression. *Biol Psychiatry* 2004; 55:390-397
38. Wiener P, Alexopoulos GS, Kakuma T, Meyers BS, Rosenthal E,

- Chester J: The limits of history-taking in geriatric depression. *Am J Geriatric Psychiatry* 1997; 5:116-125
39. Alexopoulos G: Depression in the elderly. *Lancet* 2005; 365: 1961-1970
40. Glassman AH, Occ M, Califf RM, et al: Setraline treatment of major depression in patients with acute MI or unstable angina. *JAMA* 2002; 288:701-709
41. Alexopoulos GS, Raue P, Areak P: Problem-solving therapy versus supportive therapy in geriatric major depression with executive dysfunction. *Am J Geriatric Psychiatry* 2003; 11:46-52
42. Alexopoulos GS, Kiosses DN, Klimstra S, Kalayam B, Bruce ML: Clinical presentation of the "depression-executive dysfunction syndrome" of late life. *Am J Geriatric Psychiatry* 2002; 10:98-106
43. Sneed JR, Roose SP, Sackeim HA: Vascular depression: a distinct diagnostic entity? *Biol Psychiatry* 2006; 60:1295-1298