Support for the Vascular Depression Hypothesis in Late Life Depression: Results from a Two Site Prospective Antidepressant Treatment Trial

Yvette I. Sheline, M.D., **Carl F. Pieper, Dr PH.**, **Deanna M. Barch, Ph.D.**, **Kathleen Welsh-Boehmer, Ph.D.**, **Robert C. McKinstry, M.D., Ph.D.**, **James R. MacFall, Ph.D.**, **Gina D'Angelo, Ph.D.**, **Keith S. Garcia, M.D., Ph.D.**, **Kenneth Gersing, M.D.**, **Consuelo Wilkins, M.D.**, **Warren Taylor, M.D.**, **David C. Steffens, M.D.**, **Ranga R. Krishnan, M.D.**, and **P. Murali Doraiswamy, M.D.**

Departments of Psychiatry, Radiology, Psychology and Division of Biostatistics, Washington University School of Medicine, St. Louis, MO and Department of Psychiatry, Duke University School of Medicine, Durham, NC

Abstract

Context—Research on "vascular depression" has used two approaches to subtype late life depression (LLD) based on executive dysfunction or white matter hyperintensity (WMH) severity.

Objective—Evaluate the relationship of neuropsychological performance and WMH to clinical response in LLD.

Design—2-site prospective nonrandomized controlled trial.

Setting—Outpatient clinics at Washington University and Duke University.

Participants—217 subjects age ≥ 60 met DSM-IV criteria for major depression, scored ≥ 20 (MADRS), received vascular risk factor (VRF) scores, neuropsychological testing and MRI scan; were excluded for cognitive impairment or severe medical disorders. Fazekas rating was conducted to grade WMH lesions.

Intervention—12 weeks of sertraline treatment, titrated by clinical response.

Outcome—Montgomery-Asberg Depression Rating Scale (MADRS) score over time.

Results—Baseline neuropsychological factor scores correlated negatively with baseline Fazekas scores. A mixed model examined effects of predictor variables on MADRS scores over time. Baseline episodic memory ($p = 0.002$); language ($p = 0.007$); working memory ($p = 0.01$); processing speed ($p = 0.0001$); executive function factor scores ($p = 0.002$), and categorical Fazekas ratings ($p = 0.049$) predicted MADRS scores, controlling for age, education, age of onset and race. Controlling for baseline MADRS scores these factors remained significant predictors of decrease in MADRS scores except working memory and Fazekas ratings. 33% of subjects achieved remission (MADRS ≤ 7). Remitters differed from non-remitters in baseline cognitive processing speed, executive function, language, episodic memory and VRF scores.

Corresponding Author: Yvette Sheline, M.D, Washington University School of Medicine, Department of Psychiatry, Box 8134, 660 South Euclid, St. Louis, MO 63110, Phone: 314-362-8422, Fax: 314-362-7599, yvette@npg.wustl.edu.

Discussion—Comprehensive neuropsychological function and WMH severity predicted MADRS scores prospectively over a 12 week SSRI treatment course in LLD. Baseline neuropsychological function differentiated remitters from non-remitters and predicted time to remission in a proportional hazards model. Predictor variables correlated highly with VRF severity. These data support the vascular depression hypothesis and highlight the importance of linking subtypes based on neuropsychological function and white matter integrity.

Keywords

late life depression; antidepressant; neuropsychology; WMH; cognitive deficit; age of onset; vascular risk factors; factor scores

Introduction

Late life depression (LLD) produces significant morbidity and mortality, making it a significant public health issue, given the growing number of elderly. The heterogeneity of LLD has been well described, including the large degree of medical co-morbidity, especially vascular risk factors, (e.g., cardiovascular disease, stroke, hypertension and diabetes) (1-5). Vascular disease may contribute to LLD by affecting subcortical structures involved in mood regulation and the white matter pathways that connect these structures to frontal cortex (6). Research on "vascular depression" has developed two ways of subtyping LLD: 1) those identified clinically by neuropsychological characteristics, especially executive dysfunction; and 2) those identified by brain MRI characteristics. In the subtype consisting of patients characterized as having executive dysfunction (7-8), vascular depression has been characterized clinically as a "depression-executive dysfunction syndrome of late life." (8). Despite enthusiasm for this theory (9), few studies have examined the predictive utility of cognitive function in understanding the course and outcome of LLD. A recent study by Alexopoulos et al (10) prospectively examined neuropsychological function in predicting treatment outcome in LLD and found a significant negative effect of executive function on treatment outcome, suggesting an important role for cognitive function in understanding the course of LLD.

A second subtype description of vascular depression, "MRI-defined vascular depression" (11), is defined by the presence and severity of white matter hyperintensities (WMHs), which are thought to be produced by small, silent cerebral infarctions (12). Increased WMH severity is a well-replicated finding in elderly subject groups with depression (13-19) although there are negative studies as well (20-21). Several factors contribute importantly to the pathogenesis of WMH, particularly age (22) and medical comorbidity, especially hypertension (23) diabetes mellitus (24), cardiovascular disease (29) and overall higher Framingham risk factor score (25-26).

Although each of these two ways of subtyping LLD have been shown to have clinical relevance, little work has attempted to understand the relationship between neuropsychological function and WMH in predicting course and outcome in LLD. Thus, in the current study, these two different ways of subtyping vascular depression, namely neuropsychological function variables and white matter hyperintensity (WMH) severity were used to predict course of illness over 12 weeks of antidepressant treatment in LLD. We hypothesized that impaired baseline neuropsychological performance, particularly processing speed and executive function, would predict poorer clinical response in a prospective treatment trial using sertraline. In addition, we hypothesized that increased baseline WMH severity would predict worse clinical outcome and that there would be an association between WMH and executive dysfunction in predicting poor treatment response. The study was conducted at two sites to increase sample size and our ability to generalize our results to the larger population of LLD.

Methods

Subjects

Patients were recruited for an ongoing NIMH study "Treatment Outcome in Vascular Depression" through advertising and physician referral to Washington University Medical Center and Duke University Medical Center. Of 362 phone screens at WU and 374 at Duke there were 181 clinic screens at WU and 135 at Duke (see Figure 1). Patients who met DSM-IV criteria for major depression by Structured Clinical Interview for Axis I DSM-IV Disorders (SCID-IV) administered by a research psychiatrist (YS, MD, KG or KG) were recruited into the study after satisfying exclusionary criteria. Patients were moderately depressed outpatients; no inpatients were included in the study. All patients were screened to rule out severe or unstable medical disorders (e.g. MI within past 3 months, end stage cancer, decompensated cardiac failure), known primary neurological disorders including dementia, delirium, diagnosed stroke within the past 3 months, Parkinson's Disease, brain tumors, multiple sclerosis, seizure disorder, conditions or drugs that may cause depression (e.g. systemic steroids, pancreatic cancer, uncorrected hypothyroidism), history of other Axis I disorders prior to their depression diagnosed by SCID, current suicidal risk, current episode had failed to respond to adequate trials of two prior antidepressants for at least 6 weeks at therapeutic doses, use of psychotropic prescription or nonprescription drugs or herbals (e.g. hypericum) within three weeks or 5 half lives, except for limited use of certain hypnotics or in exceptions when the patients' depression was worsening, in which case antidepressants were tapered off after starting on sertraline, or Mini Mental Status Examination score <21 (27). Patients were restricted from other therapies during participation. While our criteria excluded those with MMSE < 21, only 3% of subjects had MMSE < 24. The exclusionary criteria further reduced the patient study group to 120 enrolled patients at WU and 97 at Duke (n= 217 total). All patients were enrolled in a 12 week treatment trial with sertraline, and were restricted from other therapies during participation. At WU 109 subjects completed the protocol and 11 had early termination: 2 with side effects, 2 with psychiatric hospitalization, 1 with abnormal MRI, 4 withdrew consent, 1 non-compliant and 1had an unrelated medical illness. At Duke, 81 depressed subjects completed the protocol and there were 16 with early termination: 8 subjects with side effects, 2 with lack of efficacy, 1 unwilling to undergo the MRI, 2 withdrew consent and 3 other. Thus, there were a combined 217 intent to treat patients and a combined 190 completers. For various data analyses, there were different numbers of subjects included reflecting the partial missing data on some measures. Written informed consent approved by the relevant Institutional Review Board was obtained for all subjects. This trial is registered at clinical trials.gov Treatment Outcome of Vascular Depression NCT00045773.

Measures

Data were obtained from evaluations performed by research staff of the clinical research study at each site and included medical, psychiatric, demographic, MRI and neuropsychological measures. Demographic variables (see Table 1) were age, education gender, race, depression symptom severity (scored on the Montgomery-Asberg Depression Rating Scale-- MADRS) (28), age of depression onset, Mini-Mental status exam, final dose of sertraline and vascular risk factor (VRF) as defined by the Framingham Study (29). The Framingham Study uses a stroke risk prediction assessment tool that includes the following vascular risk factors (VRF) to predict 10-year risk of stroke in both men and women: age, systolic blood pressure, the use of antihypertensive therapy, diabetes mellitus, cigarette smoking, cardiovascular disease (coronary heart disease, cardiac failure, or intermittent claudication), atrial fibrillation, and left ventricular hypertrophy by electrocardiogram. As expected, the stroke risk increased with increasing age. In our sample subjects < age 65 had a mean VRF score of 9.0, indicating a 10 year stroke risk of 8% (average risk for the age = 7%); for those 65-74 the VRF was 12.2 indicating a 10 yr risk of 13.5% (average risk = 11%); for 75-84 the VRF was 16.2 indicating

a risk of 23% (average risk of 20%) and for 85+ the mean VRF was 19.3 with risk of 34% (average risk 13.7%). Thus, based on mean VRF scores, our population had a higher 10-year probability of stroke compared to the average stroke risk per age in the general population and the relative increase in stroke risk in our population is as follows: 14% in age group less than 65 years, 22.7% in those 65-74 years; 14.4% in those 75-84 years; and 148.1% in those aged 85 years and older.

Age at onset was ascertained from the SCID-IV and all available medical and psychiatric records. The neuropsychological testing was performed by a highly trained examiner who was supervised by a Ph.D. level psychologist (D.B. and K. W. B.). Patients were tested prior to the initiation of antidepressant medication and were psychotropic free.

Outcome Measure—Montgomery-Asberg Depression Rating Scale (MADRS) (28) scores were obtained at baseline and weekly for the 12 weeks by a research psychiatrist. Prior to study initiation a start-up meeting was held with all investigators from both sites that included training to standardize MADRS ratings across sites. For purposes of data analysis, given variable patient schedules for completing the study, completion was defined as > 8 weeks in the study. Remission was defined in patients who remained in the trial at least 8 weeks (completers) and had a final MADRS score of ≤ 7. Non-remitters were defined as patients who stayed in the trial at least 8 weeks but did not have a final MADRS score ≤ 7. While many studies have used a final MADRS score of ≤ 10 to define remission, we chose a more stringent value based on evidence from a meta-analysis (30) supporting a lower cut-off.

Sertraline Treatment—The initial sertraline dose was 25 mg for one day to rule out drug sensitivity, then 50 mg daily, with subsequent dose changes at 2 weeks (to 100 mg per day), at 4 weeks (to 150 mg), and 6 weeks (to 200 mg per day) based on treatment response and side effects. Side effects were assessed at each visit from a checklist. At any point patients who had side effects could be titrated to a lower dose. Medication adherence was assessed on each visit by self- report. Final doses and number of participants at each dose were as follows: < 100 mg-- 64, 100-125 mg—60, 150-175mg—46, 200 mg—34. Mean final dose = 114.0 (54).

Neuropsychological test performance in LLD

All participants were administered a large battery of neuropsychological tests that covered cognitive domains relevant to understanding late-life depression. We grouped the cognitive tasks into rationally motivated domains described below based on the prior literature regarding the cognitive processes tapped by each of the tasks. For further details see Sheline et al., (31). To combine the tasks, we created Z-scores for the primary dependent measure of interest at baseline across all participants and then summed the Z-scores. Follow-up waves used items 'normalized' using the mean and standard deviation at baseline. Variables in which good performance was represented by lower values, rather than higher values (such as Trails) were reverse scored to insure that higher Z-scores represented better performance for all variables. Cronbach's coefficient alpha (a measure of internal consistency) was computed for each domain.

Executive Function—This domain included verbal fluency (total phonological and semantic), Trails B (reverse scored time to completion), the color-word interference condition of the Stroop (number completed), the Initiation-Perseveration subscales of the Mattis, and categories completed from the Wisconsin Card Sorting Test. The coefficient alpha for this domain was .73.

Processing Speed—This domain included Symbol-digit modality (number completed), the color naming condition of the Stroop task (number completed), and Trails A (reverse scored time to completion). The coefficient alpha for this domain was .80.

Episodic Memory—This domain included word list learning (total correct), logical memory (total correct immediate), constructional praxis (memory performance), and the Benton Visual Retention Test (total correct). The coefficient alpha for this domain was .76.

Language Processing—This domain included the Shipley Vocabulary Test (number correct), the Boston Naming Test (number correct), and the Word reading condition of the Stroop (number completed). The coefficient alpha for this domain was .67.

Working Memory—This domain included digit span forward (number of trials correctly completed), digit span backwards (number of trials correctly completed), and ascending digits (number of trials correctly completed). The coefficient alpha for this domain was .68.

MR Imaging

MRI images were collected using a Siemens Sonata 1.5T scanner at Washington University. 3D T1-weighted (T1W) scans were acquired with MPRAGE: TR 1900 ms, TE 4ms, TI 1100ms, 222×256×128 (1×1×1.25mm). Axial T2-weighted (T2W) scans were acquired using 2D turbospin echo: TR 4000 ms, TE 97ms, 17 echoes, 2mm thickness, 10mm gap, 6 interleaves, 256×256mm, 108 slices (1×1×2mm). To improve signal-to-noise ratio (SNR), four T1W images were obtained and averaged for each subject.

MRI images were collected using a GE 1.5T scanner at Duke University. The equivalent sagittal T1W sequence was conducted using a 3D IR-prepared SPGR: TR 8.3 ms, TE 3.3 ms TI 300 ms, 256 × 256 × 124. The axial T2W scan was a 2D fast spin echo: TR = 4000 ms, TE2 = 105 ms, 5 mm thickness, field of view 150×200 mm, 20 slices (1×1×5mm). Axial FLAIR images were obtained at both sites. This T2-weighted sequence allows translation to the vast majority of clinical sites: TR = 9.99 sec; TE = 105 ms, TI = 2300 ms, slice =20, 5 mm thickness; interleaved acquisitions with no gap.

To correct for head movement and improve SNR, the four T1W scans were co-registered using standard 12 parameter affine transform to create a single average image (32). The six T2W images were collated and fused and then co-registered with the T1W scans. Both T1W and T2W images then were re-sampled to a common Talairach stereotaxic atlas (T88) using $1mm^3$ voxels. To correct for magnetic field inhomogeneities, a parametric bias field correction (PABIC) was used to correct both T1W and T2W image intensities (33-34).

T2-Weighted Hyperintensities—Hyperintensities were assessed blinded to treatment data using the modified Fazekas criteria. All ratings were conducted at Washington University School of Medicine by RCM and YIS using FLAIR and T2-weighted images together in a side-by-side review. The modified Fazekas criteria (35) describe MRI hyperintensities in three regions and follow an ascending degree of severity. The criteria assess periventricular hyperintensity (PVH) (0=absent, 1=caps, 2=smooth halo, 3=irregular and extending into deep white matter); deep white matter hyperintensities (DWMH) were scored as follows: (0=absent, 1=punctate foci, 2=beginning confluence of foci, 3=large confluent area); and subcortical gray matter lesions (SCGMH): (0=absent, 1=punctate, 2=multipunctate, 3=diffuse). Inter-rater reliability was calculated separately for the 3 Fazekas ratings: PVH (0.73); DWMH (0.86); SCGMH (0.94) and in all cases of disagreement a followup consensus rating was conducted. In addition a total Fazekas rating ("Total Fazekas score") was created by summing the 3 ratings from deep white matter, subcortical gray matter and periventricular ratings, producing a score

that ranged from 0-9. From this total score a categorical Fazekas score ("Total Fazekas categorical") was created: ≥ 3 (high) and ≤ 2 (low).

Statistical Methods

Pearson's Product Moment Correlations were used to investigate the relationship between baseline neuropsychological function and WMH (Fazekas scores). In addition, Pearson's Correlations were conducted between the Framingham Vascular Risk Factor (VRF) scores and the predictor variables.

Change in MADRS scores over 12 weeks was assessed. To accommodate missing values due to missed appointments and censoring due to dropout, a mixed model (36) was employed. Three different mixed models were then used to predict treatment outcome. Model 1: Separately for each predictor measure, neuropsychological cognitive function and WMH measures were used to predict MADRS scores following treatment controlling for time, age, education, race and age of onset (not accounting for initial MADRS). Model 2: The same predictor variables and covariates were used as in Model 1 but the model also controlled for baseline MADRS score as well as these variables in order to assess whether cognitive function or WMH predicted the magnitude of change from baseline to endpoint MADRS. Model 3: To assess the difference in trajectories, in a third analysis, the variable by time interaction was incorporated into the model to assess whether cognitive function or WMH predicted the speed of change as well as the magnitude of change from baseline to endpoint.

Prior to entering predictor variables we first examined the effect of covariates on MADRS using a mixed model to determine the results in the unadjusted model (not shown) where we only adjusted for that covariate and time. The covariates had the following bearing on the outcome: age was borderline significant ($p = 0.06$); race was not significant ($p = 0.8$); education was borderline significant ($p = 0.06$); age of onset was not significant ($p = 0.3$). These results are not displayed in Table 4. The unadjusted model for memory ($p = 0.0002$), language ($p = 0.002$), working memory ($p = 0.004$), processing speed ($p < 0.0001$), executive function ($p = 0.0004$), Fazekas cat ($p = 0.02$) indicates that all hypothesized covariates had a slightly larger magnitude effect in the unadjusted model, and, the effect slightly weakened as more covariates were added to the model, as shown in Model 1 and Model 2 in Table 4. All these hypothesized covariates were significant for the unadjusted model.

In addition, a Cox proportional hazards model (37) analyzed time to remission and was used to predict the remitter survival given baseline predictor variables and further adjusting for covariates.

Results

Demographic variables used as predictor variables of treatment outcome were age, education, race, age of onset and depression symptom severity on MADRS. The number, mean and standard deviation of these variables are shown in Table 1 and were used as covariates in the analyses. The number and percentage of patients with early onset vs. late onset depression (≥ age 60) is also shown in Table 1. MMSE scores were included and as shown in Table 1, the mean was relatively high—27.7. In addition, the mean and SD for each of the neuropsychological factor scores, the Fazekas score and the vascular risk factor score as defined by the Framingham Study (29) are shown in Table 1. Table 1 shows the demographic data for the patients who completed at least 8 weeks of the 12 week trial ("completers") vs. the patients who failed to complete at least 8 weeks ("drop-outs"). Comparing the groups, the variable that was different for drop-outs was final dose of sertraline, which was significantly lower.

Of subjects who completed at least 8 weeks of treatment, Table 2 compares those with remission of depression vs non-remitters. As shown in Table 2, 33% of all subjects achieved remission of depression ($\leq$ 7 on MADRS). The p-values indicate significant differences in variables between subjects who achieved remission ("remitters") vs. those who did not ("non-remitters"). Interestingly, compared with remitters, the non-remitters had a higher final dose of sertraline, indicating that an attempt had been made to increase the dose to a level that would achieve remission and that the difference in remission was not simply a matter of under- dosing. Figure 2 graphically displays the MADRS scores from baseline to 12 weeks of treatment for the remitters vs non-remitters.

As shown in Table 3, using a Pearson's correlation, the Framingham vascular risk factor scores were statistically significantly correlated with all of the predictor variables except for working memory. In addition, we determined correlations between the categorical Fazekas score (hi vs. low) with neuropsychological factor scores (data not shown). Fazekas scores were statistically significantly correlated with all of the baseline neuropsychological factor scores: executive function (r = -.27, p = .0005); cognitive processing (r = -.27, p = 0.0005); episodic memory (r = -.21, p = 0.004); language (r = -.15, p = 0.047); and working memory (p = 0.003).

Next, using mixed models, we examined the effect of our predictor neuropsychological measures, Fazekas scores, VRF and last sertraline dose on the trajectory of treatment response. Of note, there were different numbers of subjects in these analyses due to the different numbers of subjects completing the separate measures. We used three different prediction models for assessing the effect of baseline variables on treatment outcome. We first used a mixed model to assess the impact of predictor variables (cognitive function, Fazekas scores, VRF and last dose of sertraline) on MADRS scores, with time, age, education, age of onset and race as covariates. The following measures produced a statistically significant effect on the MADRS scores (see model 1 in Table 4): episodic memory (p = 0.002); language (p = 0.007); working memory (p = 0.01); processing speed (p = 0.0001); executive function (p = 0.002) and categorical Fazekas score (p = 0.049). In addition higher last dose of sertraline predicted worse outcome, indicating that non-remitters received a higher dose in an attempt to adequately treat their depression.

After controlling for baseline MADRS score (model 2), episodic memory (p = 0.0081), processing speed (p = 0.0006), executive function (p = .01) language (p = .03) scores, VRF scores (p = 0.03) and last dose of sertraline (p = 0.0001) all produced a statistically significant effect on MADRS scores, indicating that these variables predicted higher or lower levels of MADRS scores during the course of treatment. It was actually higher sertraline doses that predicted worse outcome. Fazekas scores and working memory scores did not significantly predict change in MADRS scores once baseline MADRS values were entered into the model.

The predictors reported in Table 4 were analyzed in separate models (in contrast to conducting a model with all predictors entered at once). As shown in Table 4, all neuropsychological factor scores had a negative relationship with MADRS score. Episodic memory, language, working memory, and executive function had a similar relationship in magnitude and direction to MADRS. The relationship between processing speed and MADRS had a larger effect than found with the other neuropsychological predictors. As the neuropsychological factor scores increased (indicating higher function) there was a decrease in the MADRS score. The relationship between total Fazekas categorical score and MADRS was positive, indicating that those with a Fazekas value of at least 3 (more severe WMH) had a larger MADRS score than with a Fazekas value less than 3. As noted above in the statistical methods section, the magnitude for all the predictors of interest (WMH and neuropsychological covariates) did slightly decrease from the model controlling for time only (not shown in Table 4) to the model controlling for time, age, age of onset, race, and education (model 1 in Table 4). When further

adjusting for baseline time and baseline MADRS (model 2 in Table 4), the magnitude of the effect for all of the neuropsychological factor scores and WMH decreased even more. This finding makes intuitive sense, since it would be expected that controlling for the baseline outcome value would account for some of the variability in the regression portion of the mixed model. The effect for working memory and total Fazekas categorical score was decreased by almost half when adjusting for baseline time and baseline MADRS, however the standard error remained about the same leading to nonsignificant results.

Neither the neuropsychological factors nor Fazekas scores interacted with time to predict MADRS scores (results not shown), controlling for age, race, education, age of onset, baseline time and baseline MADRS scores. This result indicates that while many of the neuropsychological variables predicted the overall magnitude of MADRS score change, they did not predict the rate of change (slope).

Examining the probability of remission using a Cox proportional hazards model, controlling for age, age of onset, education and race, the factors that predicted the remitter survival were episodic memory ($p = 0.006$), cognitive processing speed ($p = 0.001$) executive function ($p = 0.01$) and language function ($p < 0.05$) but not working memory or Fazekas scores.

Discussion

The principal finding of this prospective antidepressant treatment study in late life depression was that both baseline neuropsychological function, as well as white matter hyperintensity scores predicted MADRS scores over a 12 week course of treatment and that neuropsychological function and white matter hyperintensity scores were correlated. Further, all these predictor variables were highly correlated with the Framingham Vascular Risk Factor scores, indicating a strong association with vascular disease. Several studies have shown that a large number of late life depressed patients fail to respond or respond only partially to treatment, particularly those with executive impairment (7,10).

While some studies support the preeminence of executive dysfunction (9), cross-sectional assessments of cognitive function in LLD have yielded variable findings about the specificity of deficits to executive function (38-41). Of the studies using a matched control group, many (38-40) suggested the presence of disturbances across a range of cognitive domains in LLD. However, in some recent studies (41,31) disturbances occurred across a broad range of domains and could be best explained by core deficits in cognitive processing speed that influenced performance in a range of cognitive domains. Thus it is not clear whether cognitive deficits in vascular depression are specific to executive dysfunction or representative of more general disturbances in neuropsychological function that may in part reflect slowed processing speed. As noted above, compared with the number of studies examining cross-sectional neuropsychological function in LLD, there are few prospective studies of treatment outcome. The current study used a comprehensive neuropsychological battery and was thus able to simultaneously assess multiple domains of cognitive function. In the current study, even after controlling for baseline depression severity, cognitive processing speed was still strongly predictive of MADRS scores ($p = 0.0006$), whereas executive function was less highly significant ($p = 0.01$). There was also a strong predictive effect for episodic memory ($p = 0.008$). Our results add to the literature demonstrating that neuropsychological function predicts MADRS scores in LLD. They expand upon prior research by elucidating the relationship of neuropsychological function and WMH in MADRS score change. Furthermore, examining treatment remission there was a strong effect of baseline cognitive processing speed, executive function, episodic memory and language processing as well as vascular risk factor score comparing patients who achieved depression remission versus those who did not.

Similar to the effect of neuropsychological function on treatment outcome, in some studies, severity of WMH has been associated with poor antidepressant treatment response (42-44). In contrast, a study (45) that measured WMH failed to find a relationship with treatment outcome in LLD and there are clearly treatment resistant LLD subjects without vascular risk factors. However, since WMH severity was not quantified in the vast majority of treatment studies, it was not possible to compare the influence of WMH across studies. Further, very few studies have examined this question prospectively. We now add to the literature by demonstrating that WMH severity predicted MADRS scores, although not after controlling for depression severity, indicating that WMH severity was highly correlated with depression severity as well as with neuropsychological impairment. The apparently poorer performance of the Fazekas rating scale than the cognitive measures in predicting MADRS does not exclude the possibility that more sophisticated methods that include the volume of lesions and/or their location could perform better. We further note that the severity of WMH in the current study is less than in most studies that examined "MRI-defined vascular depression," however a strength of the current study is that by using a continuous rather than categorical approach we are able to examine the effect of several different predictors at the same time.

Vascular disease appears to contribute to LLD by affecting frontal white matter pathways and subcortical structures involved in mood regulation. In the current study we showed that vascular risk factors were highly correlated with both WMH and neuropsychological function, indicating that both sets of abnormalities have a vascular component. An extensive literature has provided support for the importance of WMH in LLD (11,14-19), including an effect on worsening of treatment outcome (42-44;46). There is some suggestion (34) that specific pathways are more likely to be affected in patients with vascular depression who have increased burden of WMH in those specific white matter tracts underlying brain regions important in cognition and emotion. In addition, "normal appearing" white matter may be involved in vascular depression, as manifested in diffusion tensor imaging (DTI) studies (47-49). We now demonstrate using a mixed model approach that there is a significantly worse effect of high vs. low WMH load on depression outcome. In our study, those with lowest WMH severity (total Fazekas score 0-2) on the categorical Fazekas score differed from those with higher WMH burden (Fazekas scores 3-9). It is interesting that the difference in outcome appeared to select low vs. any severity of ischemic lesion severity rather than to emphasize the more severe end of the spectrum, as was hypothesized in the concept of "MRI-defined subcortical ischemic depression" (50). It has been postulated (51) that the difference between subcortical ischemic depression and depression-executive dysfunction syndrome is the assumptions about etiology, with subcortical ischemic depression due to vascular disease and depression-executive dysfunction due to aging related changes, degenerative brain disease, an accumulation of those and other factors as well as vascular disease. In the current study the majority of patients had vascular disease as evidenced by their Framingham scores, however not sufficiently severe to cause subcortical disease, as indicated by the relatively low scores on the Fazekas subcortical gray matter index. Nonetheless, the degree of WMH did predict MADRS scores, with having some degree of WMH vs. none seeming the most important indicator. Further, we found that worse function in all neuropsychological domains was significantly correlated with Fazekas scores. Thus, in our study, there appears to be a broad involvement of neuropsychological function in predicting MADRS scores as well as in association with WMH.

An important aspect of our study is that it carefully screened for and excluded subjects with dementia, using both a MMSE cutoff of 21, clinical dementia rating (CDR) score of 0, and NINDS and DSM-IV criteria to exclude dementia. In our study the vast majority of subjects had MMSE scores of 28-30 and only 3% scored < 24. Since we have previously seen a high degree of correlation between WMH and microstructural abnormality in "normal appearing" white matter, which further correlated with neuropsychological function (49), we suggest that a fundamental aspect of vascular depression may be the disruption of normal white matter

integrity, which then results in functional deficits in neuropsychological function. Our data support the concept that the subtypes of vascular depression defined by neuropsychological function and WMH severity overlap, and that the same etiological mechanisms may account for both sets of findings. In conclusion, this study supports the importance of both the "depression-executive dysfunction syndrome of late life" as well as the "MRI-defined vascular depression" subtypes of vascular depression, suggesting that both affect treatment outcome and that they describe different aspects of the same disease. A refinement suggested by our study is that that vascular disease affects neuropsychological function more broadly than just executive dysfunction, that all WMH except the least severe have a negative impact on depression outcome and that together both deficits in neuropsychological function and severity of WMH predict worse outcome.
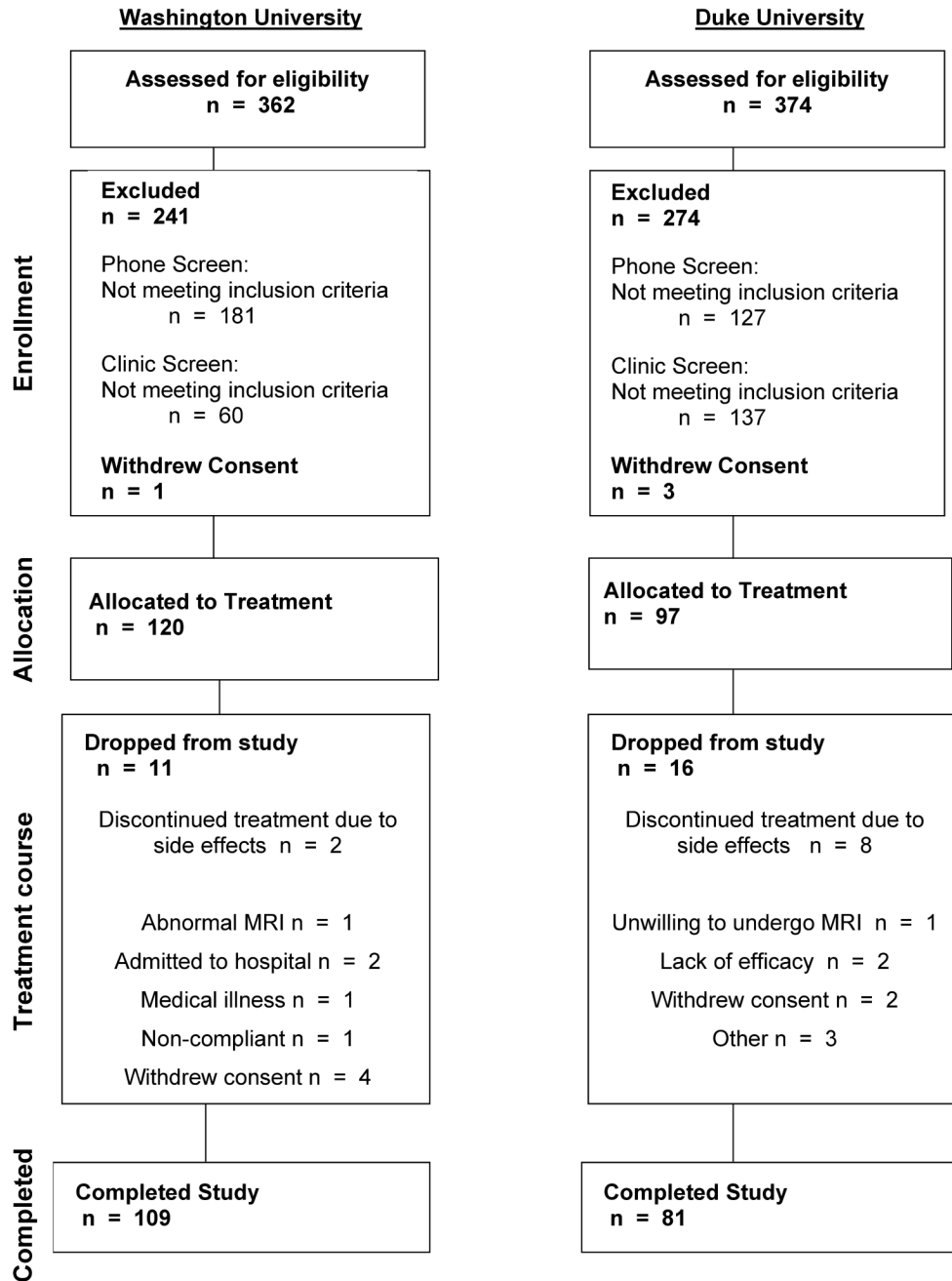
## Acknowledgments

References

1. Carney R, Blumenthal J, Stein P, Watkins L, Catellier D, Berkman L, Czajkowski S, O'Connor C, Stone P, Freedland K. Depression, heart rate variability and acute myocardial infarction. Circulation 2001;104:2024–2028. [PubMed: 11673340]

2. Glassman A, Shapiro P. Depression and the course of coronary artery disease. Am J Psychiatry 1998;155:4–11. [PubMed: 9433332]

3. Lustman P, Griffith L, Gavard J, Clouse R. Depression in adults with diabetes. Diabetes Care 1992;15:1631–1639. [PubMed: 1468296]

4. Rutledge T, Hogan B. A quantitative review of prospective evidence linking psychological factors with hypertension development. Psychosom Med 2002;64:758–766. [PubMed: 12271106]

5. Robinson R. Vascular depression and poststroke depression: where do we go from here? Am J Geriatr Psychiatry 2005;13:85–7. [PubMed: 15703316]

6. Alexander G, De Long M, Strick P. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual Review of Neuroscience 1986;9:357–381.

7. Alexopoulos G, Meyers B, Young R, et al. Arch Gen Psychiatry 1997;54:915–922. [PubMed: 9337771]

8. Alexopoulos G. New concepts for prevention and treatment of late-life depression. Am J Psychiatry 2001;158:835–838. [PubMed: 11384886]

9. Roman G. Vascular depression: An archetypal neuropsychiatric disorder. Biol Psychiatry 2006;60(12): 1306–8. [PubMed: 17157097]

10. Alexopoulos GS, Kiosses DN, Heo M, Murphy CF, Shanmugham B, Gunning-Dixon F. Executive dysfunction and the course of geriatric depression. Biol Psychiatry 2005;58:204–10. [PubMed: 16018984]

11. Krishnan R, Hays J, Blazer D. MRI-defined vascular depression. Am J Psychiatry 1997;154:497–501. [PubMed: 9090336]

12. Fujikawa T, Yamawaki S, Touhouda Y. Incidence of silent cerebral infarction in patients with major depression. Stroke 1993;24:1631–1634. [PubMed: 8236334]

13. Krishnan R, Goli V, Ellinwood E, et al. Biol Psychiatry 1988;23:519–522. [PubMed: 3345325]

14. Coffey C, Figiel G, Djang W. Leukoencephalopathy in elderly depressed patients referred for ECT. Biol Psychiatry 1988;24:143–161. [PubMed: 3390496]

15. Coffey C, Figiel G, Djang W, Weiner R. Subcortical hyperintensity on magnetic resonance imaging: a comparison of normal and depressed elderly subjects. Am J Psychiatry 1990;147(2):187–9. [PubMed: 2301657]

16. Krishnan R. Neuroanatomic substrates of depression in the elderly. J Geriatr Psychiatry Neurol 1993;6:39–58. [PubMed: 8093661]

17. Steffens D, Helms M, Krishnan R, Burke G. Cerebrovascular disease and depression symptoms in the cardiovascular health study. Stroke 1999;30:2159–2166. [PubMed: 10512922]

18. Taylor W, MacFall J, Payne M, et al. Greater MRI lesion volumes in elderly depressed subjects than in control subjects. Psychiatry Res 2005;139:1–7. [PubMed: 15927454]

19. Firbank M, Lloyd A, Ferrier N, O'Brien J. A volumetric study of MRI signal hyperintensities in late-life depression. Am J Geriatr Psychiatry 2004;12:606–12. [PubMed: 15545328]

20. Guze B, Szuba M. Leukoencephalopathy and major depression: A preliminary report. Psychiatry Res 1992;45:169–75. [PubMed: 1484908]

21. Dupont R, Jernigan T, Heindel W, et al. Magnetic resonance imaging and mood disorders. Localization of white matter and other subcortical abnormalities. Arch Gen Psychiatry 1995;52:747–55. [PubMed: 7654126]

22. Guttmann CR, Jolesz FA, Kikinis R, Killiany RJ, Moss MB, Sandor T, Albert. White matter changes with normal aging Neurology 1998;50:972–978.

23. Dufoil C, Chalmers J, Coskum O, Besancom V, Bousser M, Buillion P, MacMahon S, Mazoyer B, Neal B, Woodward M, Tamarui-Mazoyer N, Taourio C. Effects of blood pressure lowering on cerebral white matter hyperintensities in patients with stroke: the PROGRESS (Perindopril Protection Against Recurrent Stroke Study) Magnetic Resonance Imaging Substudy. Circulation 2005;112:1644–1650. [PubMed: 16145004]

24. Novak V, Last D, Alsop D, Abduljalil A, Hu K, Lepicovsky L, Cavallerano J, Lipsitz L. Cerebral blood flow velocity and periventricular white matter hyperintensities in type 2 diabetes. Diabetes Care 2006;29:1529–1534. [PubMed: 16801574]

25. Jeerakathil T, Wolf P, Beiser A, Massaro J, Seshadri S, D'Agostino R, DeCarli C. Stroke risk profile predicts white matter hyperintensity volume. The Framingham Study. Stroke 2004;35:1857–1861. [PubMed: 15218158]

26. DeCarli C, Massaro J, Harvey D, Hald J, Tullberg M, Au R, Beiser A, D'Agostino R, Wolf P. Measures of brain morphology and infarction in the Framingham heart study: establishing what is normal. Neurobiol Aging 2005;26:491–510. [PubMed: 15653178]

27. Folstein M, Robins L, Helzer J. The Mini-Mental State Examination. Arch Gen Psychiatry 1983;40:812. [PubMed: 6860082]

28. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to changes. Br J Psychiatry 1979;134:382–389. [PubMed: 444788]

29. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: A risk profile from the Framingham Study. Stroke 1991;22:312–318. [PubMed: 2003301]

30. Zimmerman M, Posternak M, Chelminski I. Defining remission on the Montgomery-Asberg Depression Rating Scale. J Clin Psychiatry 2004;65:163–168. [PubMed: 15003068]

31. Sheline Y, Barch D, Garcia K, Gersing K, Piper C, Welsh-Bohmer K, Steffens D, Doraiswamy P. Cognitive function in late life depression: relationships to depression severity, cerebrovascular risk factors and processing speed. Biol Psychiatry 2006;60:58–65. [PubMed: 16414031]

32. Buckner RL, Head D, Parker J, Fotenos AF, Marcus D, Morris JC, Snyder AZ. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. Neuroimage 2004;23:724–738. [PubMed: 15488422]

33. Styner, M.; Brechbuhler, C.; Szekely, G.; Gerig, G. IEEE Trans Med Imaging. Vol. 19. 2000. Parametric estimate of intensity inhomogeneities applied to MRI; p. 153-165.

34. Sheline YI, Vaishnavi SN, Mintun MA, Price JL, Wilkins CH, Synder AZ, Barch DM, Couture L, McKinstry RC. Regional WHM Burden in Automated Segmentation Distinguishes Late Life Depressed Subjects from Controls Matched for Vascular Risk Factors. Am J Psychiatry 2008;165:524–32. [PubMed: 18281408]

35. Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. AJR Am J Roentgenol 1987;149:351–356. [PubMed: 3496763]

36. Laird N, Ware H. Random-effects models for longitudinal data. Biometrics 1982;38:963–974. [PubMed: 7168798]

37. Klein, J.; Moeschberger, M. Survival Analysis: Techniques for Censored and Truncated Data. second. NY: Springer; 2004.

38. Kramer-Ginsberg E, Greenwald B, Krishnan R, et al. MRI signal hyperintensities in geriatric depression. Am J Psychiatry 1999;156:438–444. [PubMed: 10080561]

39. Hart R, Kwentus J. Psychomotor slowing and subcortical-type dysfunction in depression. J Neurol Neurosurg Psychiatry 1987;50:1263–1266. [PubMed: 3681304]

40. Boone K, Lesser I, Miller B, Wohl M, Berman N, Lee A, Back C. Cognitive functioning in older depressed outpatients: relationship of presence and severity of depression to neuropsychological test scores. Neuropsychology 1995;9:390–398.

41. Butters M, Whyte E, Nebes R, Begley A, Dew M. Mulsant B: The nature and determinants of neuropsychological functioning in late-life depression. Arch Gen Psychiatry 2004;61:587–595. [PubMed: 15184238]

42. Hickie I, Scott E, Mitchell P, et al. prognostic significance in patients with severe depression. Biol Psychiatry 1995;37:151–160. [PubMed: 7727623]

43. Hickie I, Scott E, Wilhelm K, et al. Subcortical hyperintensities on magnetic resonance imaging in patients with severe depression: a longitudinal evaluation. Biol Psychiatry 1997;42:367–374. [PubMed: 9276077]

44. Simpson S, Baldwin RC, Jackson A, et al. neuropsychological, and neuroradiological findings in late-life depression. Psychol Med 1998;28:1015–1026S. [PubMed: 9794009]

45. Salloway S, Boyle PA, Correia S, et al. The relationship of MRI subcortical hyperintensities to treatment response in a trial of sertraline in geriatric depressed outpatients. Am J Geriatr Psychiatry 2002;10:107–11. [PubMed: 11790641]

46. Taylor WD, Steffens DC, Krishnan KR. Psychiatric disease in the twenty-first century: The case for subcortical ischemic depression. Biol Psychiatry 2006;60:1299–1303. [PubMed: 17014829]

47. Alexopoulos G, Murphy C, Gunning-Dixon F, Latoussakis Vk, Kanellopoulos D, Klistra S, Lim K, Hoptman M. Microstructural white matter abnormalities and remission of geriatric depression. Am J Psychiatry 2008;165:236–244.

48. Murphy C, Guning-Dixon F, Hoptman M, Lim K, Ardekani B, Shields J, Hrabe J, Kanelopoulos D, Shanmugham B, Alexopoulos G. White Matter Integrity Predicts Stroop Performance in Patients with Geriatric Depression. Biol Psych 2007;61:1007–1010.

49. Shimony J, Sheline Y, D'Angelo G, Epstein A, Benzinger T, Mintun M, McKinstry R, Snyder A. Diffuse microstructural abnormalities of normal appearing white matter in late life depression: A diffusion tensor imaging study. Biol Psych. 2009 in press.

50. Krishnan K, Taylor W, McQuaid D, MacFall J, Payne M, Provenzale J, Steffens D. Clinical characteristics of magnetic resonance imaging-defined subcortical ischemic depression. Biol Psych 2004;55:390–397.

51. Alexopoulos G. The vascular depression hypothesis: 10 years later. Biol Psych 2006;60:1304–05.

Washington University

Duke University

Enrollment

| Assessed for eligibility |
| n = 362 |

| Assessed for eligibility |
| n = 374 |

Excluded
n = 241

Phone Screen:
Not meeting inclusion criteria
n = 181

Clinic Screen:
Not meeting inclusion criteria
n = 60

Withdrew Consent
n = 1

Excluded
n = 274

Phone Screen:
Not meeting inclusion criteria
n = 127

Clinic Screen:
Not meeting inclusion criteria
n = 137

Withdrew Consent
n = 3

Allocation

Allocated to Treatment
n = 120

Allocated to Treatment
n = 97

Treatment course

Dropped from study
n = 11

Discontinued treatment due to
side effects n = 2

Abnormal MRI n = 1

Admitted to hospital n = 2

Medical illness n = 1

Non-compliant n = 1

Withdrew consent n = 4

Dropped from study
n = 16

Discontinued treatment due to
side effects n = 8

Unwilling to undergo MRI n = 1

Lack of efficacy n = 2

Withdrew consent n = 2

Other n = 3

Completed

Completed Study
n = 109

Completed Study
n = 81

Figure 1.
Patient flow for Washington University and Duke University sites, indicating the numbers of subjects included in the screening process, enrollment, final allocation to the study, subjects who dropped out of treatment and reasons for discontinuation.
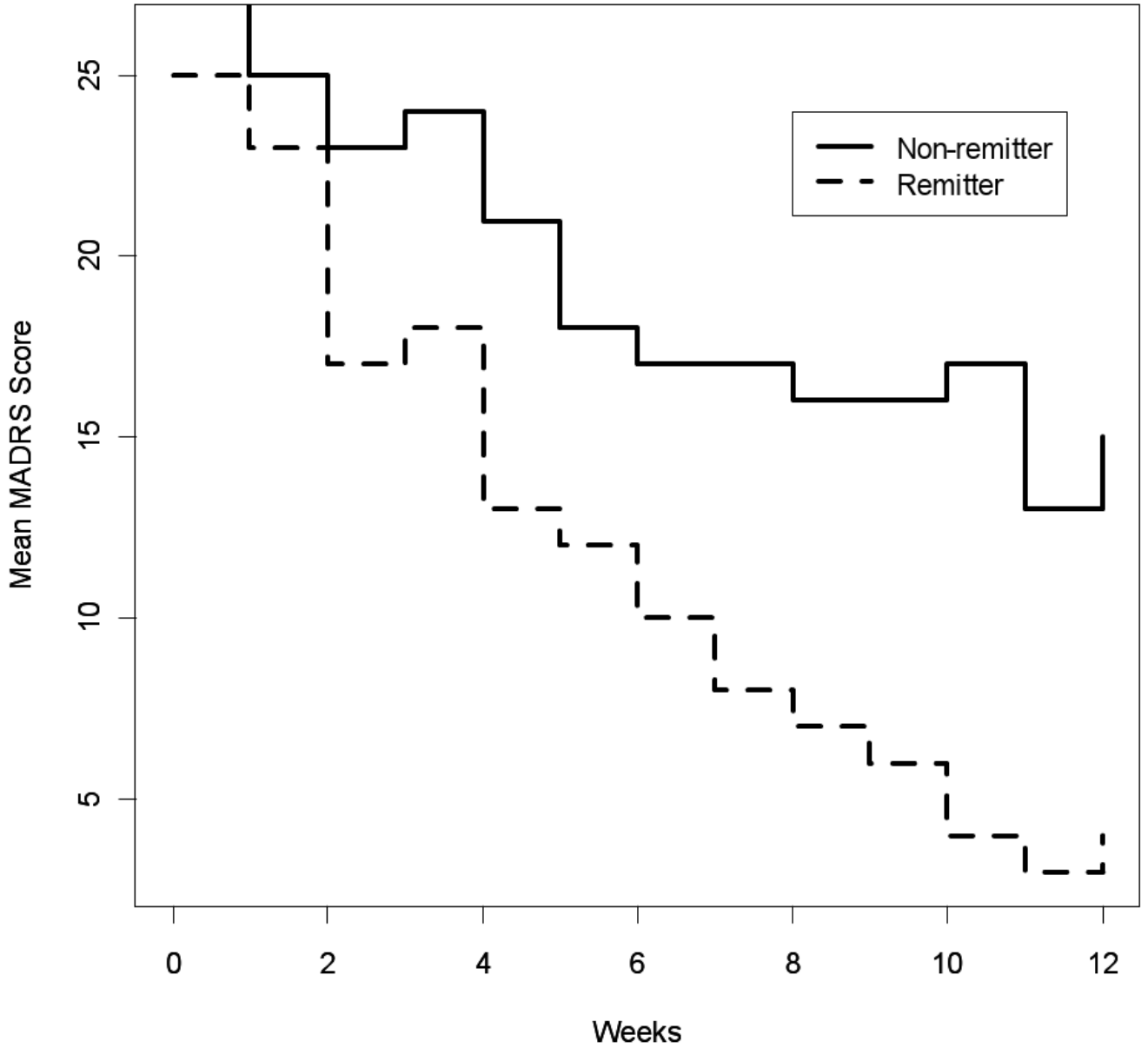
Figure 2.
The MADRS scores over the 12-week course of treatment are plotted separately for subjects
who achieved remission of depression and those who did not have remission of depression.
All subjects in this analysis remained in the study for at least 8 weeks (see text).

Table 1

Demographics, including Z scores

	Total (N=217)	Dropouts (N=27)	Completers (N=190)	p-value
White race, n (%) (n=217)	198 (91.2%)	25 (92.6%)	173 (91.1%)	1.00
Sex, male, n (%) (n=217)	96 (44.2%)	13 (48.2%)	83 (43.7%)	0.68
Late age of onset (≥ age 60), n (%) (n=199)	90 (45.2%)	7 (38.9%)	83 (45.9%)	0.63
Age, mean (SD) (n=217)	68.4 (7.2)	67.8 (6.3)	68.6 (7.3)	0.79
Age of onset, mean (SD) (n=199)	53.6 (17.2)	53.4 (15.7)	53.6 (17.3)	0.86
Education, mean (SD) (n=217)	14.2 (3.1)	13.2 (3.1)	14.4 (3.1)	0.09
MMSE, mean (SD) (n=216)	27.7 (2.0)	27.4 (2.3)	27.8 (2.0)	0.43
VRF, mean (SD) (n=211)	11.7 (4.6)	11.6 (4.6)	11.7 (4.6)	0.96
Baseline MADRS, mean (SD) (n=217)	26.0 (4.4)	25.1 (4.5)	26.1 (4.4)	0.25
Episodic Memory, mean (SD) (n=198)	-0.34 (3.1)	-0.03 (2.8)	-0.37 (3.2)	0.66
Language Processing, mean (SD) (n=193)	-0.10 (2.3)	0.12 (2.2)	-0.12 (2.4)	0.45
Working Memory, mean (SD) (n=197)	-0.04 (2.3)	-0.74 (2.1)	0.03 (2.3)	0.10
Processing Speed, mean (SD) (n=192)	0.004 (2.5)	0.54 (2.3)	-0.05 (2.6)	0.31
Executive Function, mean (SD) (n=184)	-0.10 (3.3)	0.12 (3.2)	-0.12 (3.4)	0.80
Total Fazekas Score, mean (SD) (n=184)	4.2 (2.4)	4.5 (2.5)	4.2 (2.3)	0.56
Total Fazekas Categorical (Fazekas Total>2), n (%) (n=184)	127 (69.0%)	8 (80.0%)	119 (68.4%)	0.73
Last Sertraline Dose (mg), mean (SD) (n=204)	114.0 (54.2)	72.2 (38.2)	118.0 (53.9)	**0.0007**

Note: Fisher's exact test for categorical variables and Wilcoxon Rank-Sum test for continuous variables

Table 2

Comparison of Remitters vs Non-Remitters

	Non-remitter (N=118)	Remitter (N=72)	p-value
White race, n (%) (n=190)	109 (92.4%)	64 (88.9%)	0.44
Sex, male, n (%) (n=190)	52 (44.1%)	31 (43.1%)	1.00
Late age of onset, n (%) (n=181)	51 (45.5%)	32 (46.4%)	1.00
Age, mean (SD) (n=190)	69.2 (7.7)	67.6 (6.7)	0.21
Age of onset, mean (SD) (n=181)	53.5 (17.9)	53.7 (16.6)	0.93
Education, mean (SD) (n=190)	14.2 (2.9)	14.7 (3.4)	0.19
MMSE, mean (SD) (n=189)	27.6 (1.9)	28.0 (2.1)	0.06
VRF, mean (SD) (n=188)	12.3 (4.5)	10.6 (4.6)	**0.01**
Baseline MADRS, mean (SD) (n=190)	26.6 (4.5)	25.2 (4.1)	0.06
Episodic Memory, mean (SD) (n=180)	-0.82 (3.2)	0.37 (3.1)	**0.005**
Language Processing, mean (SD) (n=175)	-0.50 (2.5)	0.47 (2.1)	**0.004**
Working Memory, mean (SD) (n=179)	-0.09 (2.3)	0.20 (2.2)	0.42
Processing Speed, mean (SD) (n=174)	-0.56 (2.8)	0.74 (1.9)	**0.0001**
Executive Function, mean (SD) (n=167)	-0.58 (3.4)	0.61 (3.2)	**0.02**
Total Fazekas Score, mean (SD) (n=174)	4.4 (2.5)	3.8 (2.1)	0.09
Categorical Fazekas (Total Fazekas >2), n (%) (n=174)	76 (72.4%)	43 (62.3%)	0.18
Last Sertraline Dose (mg), mean (SD) (n=186)	125.7 (56.9)	105.9 (46.6)	**0.02**

Note: Fisher's exact test for categorical variables and Wilcoxon Rank-Sum test for continuous variables

Table 3

Correlations between VRF and Predictor Variables

Other Predictors	VRF Pearson's correlations	
	Correlation	p-value
Episodic Memory (n=194)	-0.22	**0.002**
Language Processing (n= 189)	-0.19	**0.01**
Working Memory (n=193)	-0.12	0.11
Processing Speed (n= 188)	-0.35	**<0.0001**
Executive Function (n=181)	-0.28	**0.0002**
Fazekas Categorical Score (n=183)	0.26	**0.0004**

Table 4

Mixed Models predicting MADRS score

Predictor	Model 1 [a,b]			Model 2 [a,c]		
	Regression Coefficient (SE)	p-value		Regression Coefficient (SE)	p-value	
Episodic Memory (n=192)	-0.45 (0.14)	**0.002**		-0.35 (0.13)	**0.008**	
Language Processing (n=187)	-0.54 (0.20)	**0.007**		-0.39 (0.18)	**0.03**	
Working Memory (n=191)	-0.48 (0.19)	**0.01**		-0.25 (0.17)	0.15	
Processing Speed (n=186)	-0.72 (0.18)	**0.0001**		-0.59 (0.17)	**0.0006**	
Executive Function (n=179)	-0.45 (0.14)	**0.002**		-0.33 (0.13)	**0.01**	
Total Fazekas Cat (n=182)	1.85 (0.93)	**0.05**		1.19 (0.85)	0.16	
VRF (n=209)	0.12 (0.10)	0.25		0.20 (0.09)	**0.03**	
Last Sertraline Dose (mg) (n=192)	0.05 (0.007)	**<0.0001**		0.03 (0.007)	**<0.0001**	

[a] The effect of each predictor was analyzed separately

[b] Main effect (across all times), controlling for time, age, age of onset, race, education

[c] Main effect (across all times), controlling for time, age, age of onset, race, education, baseline time, and baseline MADRS.