Depression Duration But Not Age Predicts Hippocampal Volume Loss in Medically Healthy Women with Recurrent Major Depression

Yvette I. Sheline,1,2,3 Milan Sanghani,1 Mark A. Mintun,1,2,3 and Mokhtar H. Gado2,3

Departments of 1Psychiatry and 2Radiology 2 and the 3Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, Missouri 63110

This study takes advantage of continuing advances in the precision of magnetic resonance imaging (MRI) to quantify hippocampal volumes in a series of human subjects with a history of depression compared with controls. We sought to test the hypothesis that both age and duration of past depression would be inversely and independently correlated with hippocampal volume. A sample of 24 women ranging in age from 23 to 86 years with a history of recurrent major depression, but no medical comorbidity, and 24 case-matched controls underwent MRI scanning. Subjects with a history of depression (post-depressed) had smaller hippocampal volumes bilaterally than controls. Post-depressives also had smaller amygdala core nuclei volumes, and these volumes correlated with hippocampal volumes. In addition, post-depressives scored lower in verbal memory, a neuropsychological measure of hippocampal function, suggesting that the volume loss was related to an aspect of cognitive functioning. In contrast, there was no difference in overall brain size or general intellectual performance. Contrary to our initial hypothesis, there was no significant correlation between hippocampal volume and age in either post-depressive or control subjects, whereas there was a significant correlation with total lifetime duration of depression. This suggests that repeated stress during recurrent depressive episodes may result in cumulative hippocampal injury as reflected in volume loss.

Key words: hippocampus; depression; age; amygdala; magnetic resonance imaging; MRI; glucocorticoids; neurotoxicity; stereology
be an interaction between the two variables to produce synergistic effects in reducing hippocampal volume.

**MATERIALS AND METHODS**

Subjects: Depressed subjects were recruited to the outpatient psychiatry service at Washington University School of Medicine (WUSM), primarily from referrals from other psychiatrists who knew the inclusionary and exclusionary criteria (17/24 subjects). Control subjects were recruited primarily from the Aging and Development Project maintained by the Psychology Department at Washington University (13/24 subjects) and also from notices placed at the medical center (11/24 subjects). Subjects were recruited over a 4 year period. Subjects ranged in age from 23 to 86 years (mean = 54; median = 56). Inclusionary criteria were a history of recurrent major depression, with at least one previous episode requiring psychiatric treatment, female gender, right-handedness, and no medical illness potentially affecting the CNS. No subjects with current acute depression were included in the study to eliminate potential confounds in MRI volumetric measurements related to state-dependent changes such as hypercortisolism. None of the subjects had been acutely depressed within the past 4 months. Only women were selected to decrease the incidence of occult cardiovascular disease and to minimize gender-based brain differences (Aboitiz et al., 1992). A case-control design was used to match patients with a history of depression (post-depressed) for age and education level to controls (10 years). Groups were matched post hoc for height, a predictor of general brain size (Andreason et al., 1994). Potential subjects were screened by one of us (Y.I.S.) using a questionnaire, medical history, review of medical records, and physical exam to exclude those with medical problems potentially affecting the CNS. Exclusionary conditions included a current or past neurological disorder, head trauma, uncontrolled hypertension, myocardial infarction or ischemia, diabetes, Cushing’s disease, steroid use, and drug/alcohol abuse. All subjects had documented normotension. In addition, subjects who had received more than three courses of electroconvulsive therapy (ECT) were excluded.

Subjects gave written informed consent before inclusion in the study. Clinical assessment was conducted by a psychiatrist (Y.I.S.) experienced in the use of the Diagnostic Interview for Genetic Studies (DIGS). The DIGS is a structured interview with high reliability (Nurnberger et al., 1993) that was used to make the diagnosis of past episodes of recurrent major depression by Diagnostic and Statistical Manual of Mental Disorders-IV criteria and to exclude other psychiatric diagnoses. The DIGS was used to assess both depressed and control subjects. Current antidepressant status and dosage were determined. To determine the presence and severity of any current depressive symptoms, all patients and controls were assessed using the Hamilton Rating Scale for Depression (HAM-D).

Estimation of total time depressed. The number of symptoms and duration (days) of each episode were determined using the DIGS, as described above to define episodes of major depression. This was conducted using the Post Life Charting Method (Post et al., 1988) to anchor each episode. Cumulative duration of depression was then calculated, summing over all episodes. Rigorous diagnostic criteria were used to establish each episode, and whenever possible, corroborating information was obtained from family members or treating psychiatrists, because in depressed populations, self-reporting may underestimate the duration of earlier episodes. Patients with more severe histories of depression, such as those in the current study, have been shown to have greater stability of diagnosis (Rice et al., 1987).

Neuropsychological testing. To determine the functional significance of brain volumetric changes, neuropsychological testing was performed. Subjects were administered the following tests: Wechsler Memory Scale-Revised (WMS-R), Logical Memory I (LMI) and II (LMIII) subtests (Wechsler, 1987); Auditory Verbal Learning Test (AVLT) (Rey, 1964); Trail Making Test, parts A and B; and Wechsler Adult Intelligence Scale-Revised (WAIS-R), Information and Block Design subtests (Wechsler, 1981). All tests were administered before administration of the dexamethasone suppression test (see below).

The Logical Memory subtest was administered according to the standardized protocol outlined in the WMS-R Manual (1987). Two stories were read aloud to the subject by the examiner. The subject was asked to recall the story using a maximum of six words and then to recall it from memory after each story. One point was given for each verbatim or accepted alternate response phrase. After a 30 min delay the subject was asked to recall each story again. Subjects’ responses were recorded on audiotape and transcribed to ensure accurate scoring. Scores were tabulated for each story as well as the total score for the immediate (LMI) and delayed (LMIII) trials.

The AVLT was administered using a standard protocol (Lezak, 1995) to present a 12-word list five times followed each time by recall of as many words as possible, then a second 12-word list with recall. After a 30 min delay subjects were asked to recall the first list again. One point was given for each recalled word during each presentation. The sum of the total number of words recalled in presentations one through five was also calculated.

Trails A and B were administered in two parts according to a standard protocol (Lezak, 1995). On part A, subjects were asked to draw a line to connect consecutively numbered circles as quickly as possible, with time necessary to complete the task recorded in seconds. On part B, subjects were asked to draw a line connecting consecutively numbered and lettered circles as quickly as possible, alternating between numbers and letters (1 to A, A to 2, B to B, etc). Time to completion was recorded in seconds.

WAIS-R Information was administered according to the WAIS-R Manual (1981). Twenty-nine questions arranged in order of increasing difficulty were read directly from the manual, and the subjects’ responses were recorded. The test was discontinued after five consecutive failures. One point was scored for each correctly answered question. WAIS-Block Design was also administered according to the WAIS-R Manual (1981). Subjects were presented with designs requiring assembly of four blocks (first five designs) or nine blocks (last four designs). For each design the subject was shown a card with a picture of the design. There was a 1 min time limit on the first five designs and a 2 min time limit on the last four designs. The test was discontinued after two consecutive failed design constructions. Subjects were given points according to time recorded for completion of each correctly assembled design, and the total earned points was calculated.

Magnetic resonance imaging. MRI scans were obtained using a Magnetom SP-4000 1.5T imaging system (Siemens Medical Systems, Iselin, NJ), and a standard Siemens 30 cm circularly polarized radio frequency (rf) head coil. Anatomic images consisted of 128 contiguous 1.2-mm-thick axial slices and were acquired using magnetization-prepared rapid gradient echo (MPRAGE) acquisition. No sedation was used during scanning. Specific MPRAE scanning parameters were as follows: TR = 10 msec, TE = 4 msec, inversion time = 300 msec, flip angle = 8, matrix = 256 × 256 pixels, voxel size = 1 × 1 × 1.25 mm, slice thickness = 1.25 mm.

Image preprocessing. A graphics workstation (Sun Stargate 20, Sun Microsystems, Mountain View, CA) was used for initial image processing using ANALYZE software (Biomedical Imaging Resource, Mayo Foundation) (Robb, 1990). Images were interpolated to 0.5 mm sections. They were then reoriented to the anterior-posterior commissure plane for standard alignment. To minimize inter-scan variations, MR images underwent gray-scale normalization. Cylindrical regions-of-interest subvolumes inclusive of the amygdala and hippocampus were generated to aid in the selection of volumes. Three-dimensional MR images were preprocessed using Neurode software (Jandel Scientific, San Rafael CA), by the Marquardt-Levenberg algorithm for nonlinear curve fitting (Donald and Marquardt, 1963). The images were then scaled to create high contrast between gray and white matter in an eight-bit format.

Stereological method and reliability. Two raters (Y.I.S., M.S.) measured unilateral hippocampal gray matter volumes after training and assessment standardization with a neuroradiologist (M.H.G.) expert in hippocampal neuroanatomy. Stereological estimation methods, which have been used with precision in microscopy and MR volume determination (Mayhew, 1992) were used to estimate all hippocampal, amygdala, and whole brain volumes. Sampling parameters and grid size were set to yield at least 150 “hits” per measurement, a number that has previously been determined to yield reliable measurements in brain volume determination (Gundersen, 1988). From three-dimensional MRI cubical subvolumes composed of 0.5 × 0.5 × 0.5 mm voxels, coronal slices were sampled every 1.5 mm for hippocampus, every 1.0 mm for total amygdala, and every 0.5 mm for amygdala core nuclei. A rigid grid of points was then superimposed on the images. Grid points falling within the hippocampal, total amygdala, or amygdala core nuclei gray matter (see definitions below) were counted. Measurements were made using MEASURE software (Joines, University of Tennessee, Knoxville, TN) using orthogonal perspectives: sagittal, axial, and coronal (Fig. 1) to aid in anatomic localization. Volume estimates were calculated by MEASURE based on the number of selected grid points. Raters were blind to subject
identity and measured left and right volumes separately. Mean volumes were determined from an average of four measurements of each volume, two measurements each by two independent raters for hippocampal, amygdala core nuclei, and total cerebral volumes. Means for total amygdala volumes, left and right, were determined from an average of two independent measures by the same rater. Intra-rater and inter-rater reliability were determined using the intraclass R. Intra-rater correlation coefficients were calculated for total (0.94 and 0.95), left (0.90 and 0.95), and right (0.95 and 0.98) hippocampal volumes, total (0.97 and 0.96), left (0.94 and 0.93), and right (0.92 and 0.90) amygdala core nuclei volumes, total (0.90), left (0.86), and right (0.88) whole amygdala volumes, and whole brain volumes (0.93 and 0.93). Inter-rater correlation coefficients were calculated for total (0.95), left (0.90), and right (0.95) hippocampal volumes, total (0.89), left (0.90), and right (0.82) amygdala core nuclei volumes, and whole brain volumes (0.92). For hippocampus, 48 pairs of intra-rater measurements and 48 pairs of inter-rater comparisons were made. For amygdala, 48 pairs of intra-rater and inter-rater core amygdala measurements were made, and 48 pairs of intra-rater measurements were made for whole amygdala by one rater. For whole brain, 48 pairs of intra-rater reliability measurements were made, and 28 pairs of inter-rater reliability measurements were made.

Anatomic definition of the hippocampus. Anatomical boundaries were defined by specific rules (Duvernoy, 1988; Bartzokis et al., 1993) as follows: every third coronal slice was assessed, beginning with the posterior-most slice in which the hippocampus was visualized. Orthogonal views were consulted in cases of anatomic uncertainty, although the coronal view retained priority. Posteriorly, the tail of the hippocampus continues as the indusium griseum, a thin strip of gray matter overlying the surface of the corpus callosum. For purposes of measurement, the posterior-most slice for inclusion was defined as the slice where the hippocampus first appeared adjacent to the trigone of the lateral ventricle. Included tissues were an elongated gray matter complex bordered superiorly by the fornix-fimbria white matter junction, inferiorly by parahippocampal gyrus white matter, medially by the subarachnoid spaces of various cisterns (e.g., ambient cistern), and laterally by the CSF-filled lateral ventricle. The gray matter complex included the cornu ammonis (CA), dentate gyrus, and subiculum, (i.e., the head and body of the hippocampus were included). The vertical digitation of the head of the hippocampus, which curves up and medial to the amygdala in coronal sections, was included. Excluded tissues were the fornix-fimbria white matter complex; the alveus, the intraventricular white matter covering of the hippocampus; the white matter of the parahippocampal gyrus; vari-

Figure 1. Hippocampal volume measurement using stereology. A, Coronal section through the hippocampus. Points within the hippocampus highlighted in red were selected from a randomly placed 5 × 5 mm² grid overlying the hippocampus and are simultaneously displayed in (A) coronal, (B) sagittal, and (C) horizontal views. Arrow indicates the caudate (A), and ↔ indicates the amygdala 171 (B). D, Cubic volumes containing the hippocampus were sectioned out from the total brain volume.
The superior border of the pons was chosen as the point of demarcation because it is easily recognizable. This volume measurement was made using stereological methods as described above.

The amygdala was visualized in three planes simultaneously during measurement, and some boundaries were better visualized in one plane than in another. The noncore amygdala was defined by measuring the total amygdala and subtracting the core amygdala. The anterior boundary of the amygdala visualized in coronal section, was the first section in which the temporal stalk connected to the white matter of the insula. Dorsally, visualized in coronal section, the border was defined in anterior sections by the endorhinal sulcus between the basal forebrain and temporal lobe, and posteriorly in sagittal sections by a horizontal to the posteroinferior edge of the temporal stem with the temporal horn of the lateral ventricle.

Ventrally, visualized in sagittal section, the amygdala was bounded by a horizontal line connecting to the ventral/anterior edge of the hippocampus. Posteriorly, seen in sagittal section, the amygdala was bounded by its border with the hippocampus. Medially, seen in coronal section, the amygdala was bounded by subarachnoid space. Laterally, seen in coronal section, the amygdala was bounded by white matter. These boundaries defined the total amygdala. The noncore amygdala was defined as the total amygdala minus the core amygdala.

Total cerebral volume. All brain tissue of the cerebral hemispheres, both gray and white matter, was included in the assessment of total cerebral volume. The midbrain superior to the pons was also included. The superior border of the pons was chosen as the point of demarcation because it is easily recognizable. This volume measurement was made using stereological methods as described above.

Cortisol levels. Blood samples were obtained to measure cortisol concentrations at 8:00 A.M. and 4:00 P.M. on the day before (baseline) and the day after oral administration of 1 mg dexamethasone at 11:00 P.M. The day before (baseline) and using stereological methods as described above.

Data analysis. Two-tailed paired t tests were used to compare post-depressed and control subjects on all demographic and MRI measures.

The Pearson correlation was used to determine the significance of correlations among and between MRI measures and total cumulative duration of depression. Before using the Pearson correlation, variables were tested to determine whether they met criteria for normal distribution (Kolmogorov-Smirnov test). Pearson correlations were used to determine the correlation between age and hippocampal volumes separately in both the depressed and control samples, as well as in the total sample. A multiple regression with age and depression duration as independent variables was used to test these effects on hippocampal volumes. To examine the effect of age on hippocampal volumes, after removing the effect of depression duration, a regression analysis was conducted using age versus hippocampal volumes corrected for depression duration. This correction factor was derived from the multiple regression analysis with age and depression duration as independent variables and hippocampal volume as the dependent variable. In addition, to determine whether there was an interaction between age and depression status in predicting hippocampal volumes, these variables were combined into a new variable, and a multiple regression analysis was run using all three variables.

RESULTS

Demographic and cortisol data

Table 1 summarizes demographic and cortisol data for the 24 post-depressed subjects and their case-matched controls. The HAMD scores are consistent with the absence of acute current depression. Post-depressed subjects had a mean of 4.8 lifetime episodes of major depression (range 1–18) accounting for a lifetime mean of 1058 d depressed (range 21–3752; median, 624). Five subjects had received electroconvulsive therapy (ECT) previously during the course of their treatment and had elapsed times since last ECT treatment of 34, 30, and 14 years, and 7 and 11 months. Sixteen of the 24 post-depressed subjects were currently receiving antidepressants: selective serotonin reuptake inhibitors in four cases; tricyclic antidepressants in five cases; trazodone in two cases; and mirtazapine, nefazodone, and buproprion in one case each. Two post-depressed patients had previously been treated with neuroleptics: one with chlorpromazine and one with haloperidol. In both cases treatment duration was brief, and neither patient was psychotic at the time of treatment. Cortisol data were obtained from 17 of the total of 24 subject pairs. There was no difference between post-depressed and matched controls in either baseline cortisol levels \((p = 0.56)\) or post-dexamethasone cortisol levels \((p = 0.38)\) (Table 1), as would be expected in subjects who were not acutely depressed.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Depressed subjects ((n = 24))</th>
<th>Matched controls ((n = 24))</th>
<th>Paired t test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Age (year)</td>
<td>52.8</td>
<td>18.4</td>
<td>52.8</td>
</tr>
<tr>
<td>Education (year)</td>
<td>16.3</td>
<td>3.5</td>
<td>16.3</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>162.1</td>
<td>5.4</td>
<td>163.8</td>
</tr>
<tr>
<td>Cortisol (μg/l)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>21.86</td>
<td>11.70</td>
<td>21.37</td>
</tr>
<tr>
<td>DST</td>
<td>1.99</td>
<td>1.26</td>
<td>1.99</td>
</tr>
<tr>
<td>HAMD</td>
<td>5.9</td>
<td>4.8</td>
<td>1.3</td>
</tr>
</tbody>
</table>


\(^a\) Cortisol levels based on \(n = 17\).

\(^b\) Cortisol levels based on \(n = 17\).

\(^c\) Calculated from 14 matched pairs.
In this sample, 26 women were postmenopausal and 22 were premenopausal. Of the 26 postmenopausal women, 14 were from the post-depressed group and 12 were controls ($\chi^2$ p value = 0.67). In the total sample, 17 women were on hormone replacement therapy. Of women on hormone replacement therapy, 9/17 were post-depressed and 8/17 were controls ($\chi^2$ p value = 0.85).

**Brain volumetric data**

Table 2 summarizes brain volumetric data for post-depressed subjects and matched controls. Mean total, left, and right hippocampal gray matter volumes were significantly smaller for post-depressed subjects than for controls. On average the differences were 9, 10, and 8% for the total, left, and right volumes, respectively (Fig. 2). Mean total, left, and right amygdala core nuclei volumes were significantly smaller in post-depressives than in control subjects; on average 13, 13, and 12% for total, left, and right volumes, respectively (Table 2), whereas amygdala noncore volumes, total amygdala volumes, and total cerebral volumes did not differ. Examining hormone replacement therapy alone as a predictor of total hippocampal volume, there was no statistically significant difference in volume between subjects on hormone replacement therapy and those not on hormone replacement therapy ($t$ = -0.08, $p$ = 0.94).
replacement therapy (4699 mm$^3$) and those not on therapy (4780 mm$^3$) ($p = 0.67$). Postmenopausal status alone did not predict total hippocampal volume (4845 mm$^3$ premenopausal vs 4620 mm$^3$ postmenopausal; $p = 0.23$).

Correlations between hippocampal gray matter volumes, age, menopausal status, and total time depressed

In the total sample, the correlation between age and total hippocampal volume was $r = -0.19$ ($p = 0.20$). For post-depressed subjects the correlation between total hippocampal volume and age was $r = -0.32$ ($p = 0.13$). In control subjects the correlation between age and total gray matter hippocampal volume was $r = -0.08$ ($p = 0.71$). Analysis of interaction indicated that there was no interaction between age and depression status ($p = 0.43$). In contrast, there was a significant inverse correlation between total days depressed and total hippocampal gray matter volume ($r = -0.60; p = 0.002$) (Fig. 3). This was true for both left hippocampal gray matter volume ($r = -0.57; p = 0.003$) and right hippocampal gray matter volume ($r = -0.55; p = 0.005$).

To correct for a contribution of depression duration to the relationship between age and hippocampal volume, a regression analysis was conducted using age versus hippocampal volume corrected for depression duration. This analysis showed no significant correlation with age ($r = -0.17, p = 0.26$) (Fig. 4).

A multiple regression analysis did not find a significant relationship ($p = 0.61$) using hormone replacement therapy, postmenopausal status, and age as independent variables and total hippocampal volume as the dependent variable. Examination of left and right hippocampal volumes separately also did not yield significant relationships.

Correlations between hippocampal gray matter volumes and amygdala core nuclei volumes

The correlation between total hippocampal gray matter volume and total amygdala core nuclei volume for post-depressed subjects was $r = 0.68, p = 0.0001$ (Fig. 5), and this held for the left ($r = 0.65, p = 0.0004$) and right sides ($r = 0.58, p = 0.002$) as well. In control subjects these respective correlations were $r = 0.50 (p = 0.01), r = 0.54 (p = 0.006)$, and $r = 0.42 (p = 0.04)$. In contrast, the correlations in post-depressed subjects between total and both left and right hippocampal gray matter volumes and total ($r = -0.11, p = 0.63$) (Fig. 5), left ($r = 0.15, p = 0.49$), and right ($r = -0.02, p = 0.93$) amygdala noncore nuclei volumes, respectively, were not significant. The respective correlations in control subjects ($r = 0.46, p = 0.02; r = 0.45, p = 0.03$; and $r = 0.42, p = 0.04$) with amygdala noncore nuclei volumes were significant.

Neuropsychological data

As shown in Table 3, scores for list learning, total recall and list learning, long delay were significantly decreased in subjects with a history of depression compared with controls. Paragraph recall scores did not achieve significance. Trails A, Trails B, and WAIS information subscale scores were not significantly different between the groups. The WAIS block design subscale score was significantly less in subjects with a history of depression than in control subjects.

Effect of removing subjects who had received ECT

The data analyses comparing hippocampal gray matter volumes between post-depressed and control subjects were rerun after removal of the five subjects who had received ECT and their matched controls. There was still a significant difference in total ($t = -2.1, p = 0.04$) and left ($t = -2.4, p = 0.03$) hippocampal gray matter volumes and a trend toward a difference in right ($t = -1.8, p = 0.08$) hippocampal volumes between post-depressed and control subjects. In post-depressed subjects there was also a significant correlation between days depressed and both left ($r = -0.59, p = 0.005$) and right ($r = -0.53, p = 0.01$) hippocampal volumes.

DISCUSSION

The main findings of this study are the lack of a correlation between age and hippocampal volume in medically healthy subjects and the finding of smaller hippocampal volumes in subjects with a history of depression. These subjects also had evidence of poorer verbal memory performance, despite being in remission. Extending our pilot study (Sheline et al., 1996), we report that 24 women with a history of depression, but no medical comorbidity, had bilaterally smaller hippocampal volumes than age-matched controls. Findings of decreased hippocampal volumes have been reported in acutely depressed subjects (Narayan et al., 1998) and in patients suffering from post-traumatic stress disorder (Bremner et al., 1995). We show that post-depressed women exhibited...
abnormalities on a neuropsychological measure of hippocampal function, the AVLT, suggesting that their reduced hippocampal volumes were the result of functionally significant damage. Furthermore, hippocampal volumes correlated inversely with the cumulative lifetime duration of depression. In contrast to differences in hippocampal volumes and AVLT scores, post-depressed patients did not differ from control subjects in overall brain size or on general intellectual performance as assessed by the WAIS information subscale.

Given age- and depression-related increases in medical conditions, it is critical to know whether patients with changes in brain structure volumes had concurrent physical illnesses that could enhance neuronal vulnerability. In the present study, unlike most previous studies, we selected only subjects with no current or past medical or neurological conditions. An increased incidence of microinfarction has been linked to both chronic hypertension (Kobayashi et al., 1991) and diabetes (Aronson, 1973; Desmond et al., 1993). We excluded patients with any history of substance abuse because alcohol dependence can cause cerebral atrophy (Charness, 1993). Patients with a history of myocardial infarction were excluded because it is a cerebrovascular disease risk factor and produces a large independent increase in the incidence of depression (Carney et al., 1987).

Contrary to our initial hypothesis, there was not a significant age-related hippocampal volume loss in either post-depressed or control subjects. Some previous studies of hippocampal volume in “normal” aging found a decrease starting at age 50 (Jernigan et al., 1991), whereas others did not (Lim et al., 1990). Our sample showed slight (~1% per decade) nonsignificant decreases in hippocampal volumes with age, more so in the depressed population, although there was no interaction between age and depression status. It is possible that aging per se may not lead to neurodegeneration in the hippocampus but rather may increase vulnerability to other factors causing cell death, and these factors may have differed across study populations. It is possible that by selecting subjects without medical comorbidity we may have created a sample of “supernormals” who do not show a decrease in brain structure volumes with age.

One confounding variable is ECT. Although there has been no direct evidence demonstrating ECT-induced structural brain changes (Devanand et al., 1994), animal electroconvulsive stimulation (ECS) studies suggest that brief kindled seizures may induce selective hippocampal neuronal loss (Cavazos et al., 1994). ECS exposure sufficient to induce kindling, however, far exceeds the human equivalent ECT electrical doses. We did not exclude ECT-treated patients, because they had the most severe histories of depression. A post hoc analysis after removing the five ECT-treated subjects showed a significant difference in total hippocampal gray matter volume (p = 0.04) and a correlation between total days depressed and total hippocampal gray matter volume.
volume ($p = 0.004$). Another confound was ongoing antidepressant therapy in most post-depressed subjects, although there is no known evidence that therapeutic levels are neurotoxic or that antidepressant exposure reduces brain volumes.

Additional potentially confounding variables, which we examined post hoc, were menopausal status and hormone replacement therapy. The effects of estrogen on rat hippocampus dendritic spine density (Wooley et al., 1997) and on neuropsychological function in postmenopausal women in some studies (Resnick et al., 1998) make this an important variable to examine. We did not find any significant differences between the groups in either postmenopausal or hormone replacement status, making it unlikely that these factors could account for differences in brain volumes. When we examined the issue of whether there were differences in hippocampal volume between postmenopausal and premenopausal subjects and between subjects who had received hormone replacement therapy versus those who had not, we did not find significant differences. However, we have a small sample size, and it would be interesting to examine this question in a larger study prospectively with concurrent serum hormone level determination.

Our study found significant differences between post-depressed subjects and controls in a neuropsychological measure of hippocampal function, AVLT list learning, total score, and delayed score. In subjects with more severe hippocampal damage caused by mesial temporal sclerosis, the degree of left hippocampal atrophy in patients with left temporal lobe epilepsy was correlated with severity of verbal memory deficits as measured by AVLT total and delayed recall (Kilpatrick et al., 1997). In addition, in nondemented subjects who developed Alzheimer's disease during longitudinal followup (Tierney et al., 1996), the single biggest neuropsychological predictor was the AVLT score. Subjects in both groups had known hippocampal pathology that correlated with abnormalities in the AVLT. Our finding of abnormalities in AVLT in post-depressed subjects who also had hippocampal volume decreases appears to indicate hippocampal dysfunction. The groups did not differ, however, in logical memory test performance, another measure of hippocampal function (Squire, 1992). Depressed subjects did have lower WAIS block design subscale scores. Bremner et al. (1995) found that post-traumatic stress disorder patients performed more poorly on the Benton Visual Retention task, a test that reflects spatial memory and depends on right hippocampal integrity. Another possibility is that post-depressed subjects performed more poorly on timed tasks. However, both the WAIS block design and Trails A and B are timed tests; there was no difference in performance on Trails A and B between post-depressed and control subjects.

Although further study will be needed to fully elucidate the nature of the observed association between recurrent depression and volume reduction, evidence suggests that the link may be causal—that sustained depression in humans may cause hippocampal damage, resulting in a reduction of tissue volume and memory function. A limitation of the current study is that reliability of the Post Life Charting Method in determining retrospective measures of depression duration has not been established. There is no reason to believe that the data create a biased relationship between depression duration and hippocampal volumes, however, because volume data were obtained independently and blindly. The quantitative correlation between depression duration and hippocampal volume is explained most directly by such a causal relationship in that progressively more severe changes occur in the hippocampus as the length of time depressed, and presumably the length of time with abnormal GC levels, increases. Furthermore, in animals, repeated episodes of stress have been shown to damage the hippocampus. This appears to be mediated by GC-induced neurotoxicity (for review, see Sapolsky et al., 1986; Reagan and McEwen, 1997). In addition, there is evidence in normal elderly humans that long-term exposure to GCs predicts hippocampal atrophy and memory deficits (Lupien et al., 1998).

The relevance to depression of these studies showing that chronically elevated GC levels damage hippocampal neurons depends, however, on the assumption that depression is associated with dysregulation of the glucocorticoid system. Almost 40 years ago, excessive cortisol secretion during depression was first reported (Gibbons and McHugh, 1962). Hypercortisolism and insensitivity to feedback suppression during depression have been extensively investigated to determine the relative contributions of adrenal hypersensitivity to ACTH (Amsterdarm et al., 1989), pituitary resistance to GC feedback (Holsboer et al., 1987), pituitary hypersensitivity to CRF and other hormones (Gold et al., 1984), and resistance at the hippocampus (Sapolsky et al., 1991; Young et al., 1991). Although we do not have cortisol levels from previous depressive episodes, many of our subjects had histories of hospitalization, ECT treatment, and multiple episodes, indicating relatively more severe major depression, a predictor of elevated cortisol and dexamethasone suppression test nonsuppression (Whiteford et al., 1987). Taken together, these findings suggest that HPA axis dysregulation in depression can produce repeated episodes of hypercortisolemia, which may result in hippocampal neurotoxicity.

The precise mechanisms whereby chronic glucocorticoid exposure leads to hippocampal cell death are not defined, but enhanced vulnerability to excitotoxicity may be a main factor (Armanini et al., 1990). We have shown a reduction in volumes of the core nuclei of the amygdala in recurrent major depression (Sheline et al., 1998) and now demonstrate a high correlation between hippocampal and amygdala core nuclei volume loss. Because glutamatergic pyramidal cells in the core nuclei of the amygdala are predominant and reciprocally connected with the hippocampus, it may be that overexcitation in one part of this circuit produces damage in connected structures. Furthermore, there was no difference in volumes of noncore amygdala nuclei, which contain primarily nonglutamatergic cells (Price et al., 1987).

In addition to glucocorticoid-induced neurotoxicity, other mechanisms consistent with hippocampal volume loss in recurrent depression include neurotoxic effects of corticotrophin-releasing factor (CRF) and a decrease in trophic factors such as brain-derived neurotrophic factor (BDNF). Elevated CRF levels have been described in both depressed humans (Nemeroff et al., 1984) and animal models of depression and have neurotoxic effects (Wong et al., 1995). Furthermore, CRF antagonists are neuroprotective against seizure-induced excitotoxicity (Maeceker et al., 1997). Stress decreases the expression of BDNF (Duman et al., 1997), and BDNF is neuroprotective (Mamounas et al., 1995). These mechanisms—glucocorticoid toxicity, CRF neurotoxicity, and BDNF loss resulting in increased neuronal vulnerability—are not mutually exclusive and could all be operating to alter the vulnerability of hippocampal neurons in depression.

In summary, the present study shows an association between depression and structural changes in the hippocampus. It should be pointed out that volume changes cannot be equated with neuropathological studies directly performing cell counts; only the latter would provide direct evidence for cell loss. This gray
matter hippocampal volume loss does not appear to be related to aging but could be related to factors commonly associated with aging, such as medical illnesses, which were not examined in this study. Depression-related volume loss does appear to be cumulative, suggesting that immediate recognition and treatment of depressive episodes is important in preventing cumulative damage that occurs with repeated episodes of depression.

REFERENCES
McIntosh LJ, Sapolsky RM (1996) Glucocorticoids increase the accumulation of reactive oxygen species and enhance adriamycin-induced toxicity in neuronal culture. Exp Neurol 141:201–206.
brain and the mechanisms of neuron death, pp. 305–339. Cambridge, MA: MIT.


