

## Differences in brain volumes among males and female hormone-therapy users and nonusers

Daniel L. Greenberg <sup>a,d,\*</sup>, Martha E. Payne <sup>b,d</sup>, James R. MacFall <sup>c,d</sup>,  
James M. Provenza <sup>c</sup>, David C. Steffens <sup>b</sup>, Ranga R. Krishnan <sup>b</sup>

<sup>a</sup>*Department of Psychology, UCLA, Los Angeles, United States*

<sup>b</sup>*Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, United States*

<sup>c</sup>*Department of Radiology, Duke University Medical Center, Durham, NC, United States*

<sup>d</sup>*Neuropsychiatric Imaging Research Laboratory, Duke University Medical Center, Durham, NC, United States*

Received 17 May 2005; received in revised form 3 September 2005; accepted 8 January 2006

### Abstract

Numerous studies have shown gender differences in the brain volumes of elderly adults. Some evidence shows that higher estrogen levels may be neuroprotective, suggesting that hormone therapy (HT) may in part be responsible for these gender differences; however, few studies have examined the relation between HT and brain volumes. Brain volumes of caudate, putamen, hippocampus, gray matter, white matter, white-matter lesions, and cerebrospinal fluid were measured on magnetic resonance imaging scans. A comprehensive neuropsychological battery was administered. Women were separated into two groups based on HT use, and we used multiple regression analyses to compare these groups with one another and with men. Results of brain-volume measurements showed that HT users had significantly less gray matter and more cerebrospinal fluid than nonusers. Results of the neuropsychological testing showed that HT users performed better on the Shipley Vocabulary Test than males did.

© 2006 Elsevier Ireland Ltd. All rights reserved.

**Keywords:** Volumetry; Hormone-replacement therapy; MRI; Estrogen

### 1. Introduction

Substantial research indicates that the brains of men and women age differently. In general, atrophy of the male brain begins at an earlier age and progresses more rapidly (Gur et al., 1991; Golomb et al., 1993; Cowell et al., 1994; Raz et al., 1995; Murphy et al., 1996; Xu et al., 2000; Greenberg et al., submitted for publication), although there are exceptions to this finding, both in

cross-sectional (Resnick et al., 2000) and longitudinal (Resnick et al., 2003) studies. The female hippocampus, however, is smaller (Raz et al., 1997) and shrinks more rapidly (Murphy et al., 1996; Gur et al., 2002; but see Golomb et al., 1993 for a contradictory finding). Similarly, caudate atrophy is more severe in the left hemisphere in men, but in the right hemisphere in women (Gunning-Dixon et al., 1998). Other studies indicate that atrophy may be symmetrical in women, but more severe in the left hemisphere in men (Gur et al., 1991; Cowell et al., 1994; but see Xu et al., 2000 for a contrary finding).

The etiology of these sexual dimorphisms remains unclear, although there are probably multiple causes. One

\* Corresponding author. UCLA Psychology Department, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095, United States. Tel.: +1 818 926 0421; fax: +1 310 206 5895.

E-mail address: [dlgreenberg@mednet.ucla.edu](mailto:dlgreenberg@mednet.ucla.edu) (D.L. Greenberg).

line of evidence suggests that hormone therapy (HT) may play a role. HT was once thought to reduce the severity of several neurological conditions, including dementia (Mortel and Meyer, 1995; Costa et al., 1999; Panidis et al., 2001), and the effects of stroke (Hurn and Macrae, 2000), but there have been many inconsistent findings (Shaywitz and Shaywitz, 2000; Paganini-Hill, 2001; Women's Health Initiative Writing Group, 2002). Thus, it seems possible that HT could affect brain volumes, but only a few studies have addressed this question. One study found that HT had no noticeable effect on lesion volumes, although HT users did have larger ventricles (Luoto et al., 2000); other studies found that HT users had fewer clinical abnormalities than nonusers (Schmidt et al., 1996; Cook et al., 2002). One preliminary report indicated that HT users had larger hippocampi (Eberling et al., 2003). We therefore decided to examine the relations between HT and brain volumes in a sample of healthy elderly adults. Based on the claim that HT has a protective effect, we predicted that female HT users would have larger corrected brain volumes than female nonusers or men.

## 2. Methods

The study population was recruited from a pool of normal elderly used as controls for the Conte Center for the Neuroscience of Depression at Duke University. Eligibility for this study was restricted to those aged 60 years or older who could speak and write English. Exclusion criteria included (1) major psychiatric illness; (2) active alcohol or drug dependence; (3) primary neurological illness, such as dementia, stroke, Parkinson's disease, seizure disorder and multiple sclerosis; (4) medications or medical illness that may affect cognitive function; (5) physical disability which precludes cognitive testing; and (6) metal in the body which precludes magnetic resonance imaging (MRI). Study psychiatrists verified that all patients who screened positive for depression met criteria for major depression and did not have another major psychiatric disorder such as schizophrenia, schizoaffective disorder, bipolar disorder and dementia.

The purpose of the Conte Center and its procedures were explained to each subject, and those who provided written informed consent were enrolled. Drug use and medical history were collected via interview. Participants were asked if they had any history of asthma, diabetes, heart trouble, hypertension, arthritis, stroke, cancer, emphysema, ulcers, atherosclerosis, neurological problems, anemia, or other chronic health conditions.

Participants also underwent a battery of neuropsychological tests. This battery was composed of the Consortium to Establish a Registry in Alzheimer's Disease

(CERAD) neuropsychological battery (Morris et al., 1989), a collection of neuropsychological measures with normative standards for the elderly and established utility in longitudinal studies of cognitive impairment (Welsh et al., 1994). The CERAD measures include (1) the MMSE; (2) language tasks consisting of category fluency (animal naming) and object naming (Kaplan, Goodglass, and Weintraub, 1983); (3) constructional praxis and visual memory, requiring copy of four geometric designs, with delayed recall and delayed recognition procedures; and (4) verbal learning and memory, consisting of immediate recall of three learning trials of a 10-item word list, delayed recall of the list, and recognition/discrimination of target words from nontarget foils. The CERAD battery is supplemented by other common neuropsychological measures used in clinical practice for assessing (1) immediate and delayed verbal memory (Logical Memory subtest of the Wechsler Memory Scale-Revised; Wechsler, 1987), (2) visual immediate memory (Benton Visual Retention Test; Benton, 1974), (3) verbal initiation/lexical fluency (Controlled Oral Word Association Test from the Multilingual Aphasia Examination; Benton et al., 1983), (4) attentional/executive functions [Trail Making Test (Reitan, 1992); Symbol Digit Modalities Test (Smith, 1982), Digit Span subtest of the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), and a separate ascending Digit Span task modeled after the Digit Ordering Test (Hoppe et al., 2000)], and (5) premorbid verbal ability (Shipley Vocabulary Test; Zachary, 1991). In addition, at the time of testing, a knowledgeable informant completed the Dementia Severity Rating Scale (Clark and Ewbank, 1996), a measure of cognitive and functional status. This neuropsychological assessment battery has proved useful in our previous work in depressed cohorts, allows accurate case ascertainment of dementia in large-scale epidemiological studies (e. g. Breitner et al., 1994) and is effective in tracking the longitudinal changes of Alzheimer's disease (AD) (Breitner et al., 1994). It also meets six important criteria we were striving to achieve: (1) it is well tolerated by the elderly, (2) it includes measures sensitive to early changes of AD (Robinson-Whelen and Storandt, 1992), (3) it is reasonably comprehensive, (4) it captures a range of cognitive abilities (minimizes floor and ceiling effects), (5) it is sufficiently "mainstream" that results can be easily interpreted by most researchers in dementia, and (6) it is sufficiently simple to allow reliable administration by trained technicians. The neuropsychological assessment requires approximately 60 min to administer and has been well tolerated by the vast majority of

subjects. To minimize possible fatigue effects, subjects received a 5-min rest period after 20 min of testing.

All enrolled subjects underwent a magnetic resonance imaging scan of the brain. Subjects were imaged under an Institutional Review Board Approved protocol, with a 1.5 T whole-body MRI system (Signa, GE Medical Systems, Milwaukee, WI) using the standard head (volumetric) radiofrequency coil. Padding was used to immobilize the head without causing discomfort. The scanner alignment light was used to adjust the head tilt and rotation so that the axial plane lights passed across the cantho-meatal line and the sagittal lights were aligned with the center of the nose. A rapid sagittal localizer scan was acquired to confirm the alignment.

The first set of images was obtained with an axial, multisection, T1-weighted pulse sequence (TR=500 ms, TE=15 ms) with a 256×192 data acquisition matrix, 5-mm section thickness, a 20-cm field-of-view (FOV), 1 excitation per phase-encoding increment (1 Nex) and a 32 kHz (±16 kHz) full imaging bandwidth. This was followed by a long TR (2500 ms), double-echo (TE=30 and 80 ms) spin-echo data-acquisition sequence using the same FOV, section thickness, bandwidth and spacing, 256×192 data acquisition matrix, and 1 Nex. Saturation of spins outside the imaging volume (standard gap 15 mm) and flow compensation (gradient moment nulling) was employed to eliminate artifacts due to flowing blood and cerebrospinal fluid. These images were obtained in two separate acquisitions with a 5-mm gap between sections for each acquisition. The second acquisition was offset by 5 mm from the first so that the resulting data set consisted of contiguous sections.

Two sets of dual-echo fast spin-echo (FSE) acquisitions were obtained: one in the axial plane for morphometry of most cerebral structures and (for a subset of patients) a second in a coronal oblique plane for morphometry of the hippocampus. The pulse sequence

Table 2  
Demographics by HT group

	Men	Non-HT women	HT women	<i>F</i> (2,122)	<i>P</i>
<i>N</i>	33	51	41		
Age (mean±S.D.)	71±7	71±6	70±6	0.70	0.50
Education (mean±S.D.)	16±2	15±2	15±2	1.42	0.24
Race ( <i>n</i> )					
White	24	38	40		
Black	8	12	1		
Other/unknown	1	1	0		
Hippocampus ( <i>n</i> )	10	32	31		
All other volumes ( <i>n</i> )	33	51	41		

parameters were TR=4000 ms, TE=30, 135 ms, 32 kHz (±16 kHz) full imaging bandwidth, echo train length=16, a 256×256 matrix, 3-mm section thickness, 1 Nex and a 20-cm FOV. The images were acquired in two separate acquisitions with a 3-mm gap between sections for each acquisition. The second acquisition was offset by 3 mm from the first so that the resulting data set consisted of contiguous sections. For the near coronal acquisition, the localizer scan was used to identify the anterior commissure–posterior commissure (AC–PC) line. Oblique, near-coronal images were prescribed perpendicular to this line, covering the entire brain from just anterior of the temporal lobe to a plane posterior to the lateral ventricles.

Putamenal and hippocampal volumes were determined using the GRID Program, which was developed in-house. All other volume measurements used a locally modified version of MrX Software, which was created by GE Corporate Research and Development (Schenectady, NY) and originally modified by Brigham and Women's Hospital for image segmentation (Boston, MA). The segmentation protocol used was a modified version of that developed by [Kikinis et al. \(1992\)](#) and [Byrum et al. \(1996\)](#). This supervised, semi-automated method used the multiple MR contrasts available to identify different tissue classifications through a 'seed-ing' process wherein a trained analyst manually selected pixels in each tissue type that was to be identified (gray matter, white matter, CSF, lesions (both gray and white), and background) ([Payne et al., 2002](#)). First, the brain was segmented into tissue types; the non-brain tissue, cerebellum, and brainstem were stripped away through a masking procedure, and the resulting volume was taken as the intracranial volume (ICV). Specific regions of interest (ROI) were assessed using tracing and connectivity functions. The final step was to run a summarizing program that calculated the volume of each tissue type

Table 1  
Reliability measures

Structure	ICC
Cerebrum	0.997
Gray matter	0.8
White matter	0.8
Ventricular CSF	0.9
Nonventricular CSF	0.9
White-matter lesions	0.991
Left caudate	0.9
Right caudate	0.9
Left putamen	0.8
Right putamen	0.7
Left hippocampus	0.8
Right hippocampus	0.9

Table 3  
Prevalence of comorbid conditions

	Males ( <i>n</i> =33)	Non-HT females ( <i>n</i> =51)	HT females ( <i>n</i> =41)
Asthma	1	2	1
Diabetes	4	5	0
Cardiac disease	1	6	3
Hypertension	12	9	7
Arthritis	15	31	21
Stroke	0	0	0
Cancer	0	4	0
Emphysema	0	4	0
Ulcers	0	0	0
Atherosclerosis	1	0	0
Neurological	0	0	0
Anemia	1	1	3
Other	15	17	9

within the specific ROI defined by the analyst. The GRID Program allowed for semi-automated determination of ROI volumes and was based upon a manual point-counting method (MacFall et al., 1994; Steffens et al., 2000). A combination of tissue classification with manual tracing allowed for rapid determination of ROI volumes.

All image analysis technicians received extensive training by experienced analysts. Reliability was established by repeated measurements on multiple MR scans before raters were approved to process study data. Intraclass correlation coefficients are presented in Table 1. We conducted regression analyses using SAS 8.1 (SAS Institute, Cary, NC). In brain-volume analyses, we constructed separate models for each ROI volume; to adjust for multiple comparisons, we applied the Bonferroni correction and set our alpha level at 0.002. We recentered the age and education variables about their

respective means to avoid problems with collinearity. Researchers disagree about the best way to control for differences in body size; some advocate using ICV as a covariate, while others suggest using height. In our view, using ICV as a covariate yields the clearest and most meaningful results when conducting analyses across genders (see Greenberg et al., submitted for publication, for a discussion of this issue).

### 3. Results

The study included 33 male participants. We divided female participants into groups based on their use of HT. We began with 122 female participants; 10 women were excluded because they had taken HT briefly (mean  $\pm$  S.D. =  $3.2 \pm 2$  years) in the past but were no longer taking it, and 10 women were excluded because they were current HT users who had been taking HT for less than 10 years (mean  $\pm$  S.D. =  $5 \pm 2$  years). Fifty-one women composed the non-HT group; they had never used HT. Forty-one women formed the HT group; they had been using HT for at least 10 years (mean  $\pm$  S.D. =  $23 \pm 7$  years) and were still taking HT at the time they were scanned. We chose this cutoff because previous research had suggested that current HT users only experienced a reduced risk of AD if they had been taking HT for more than 10 years (Zandi et al., 2002).

We were able to collect drug and dose information on a subset of women in the HT group. Drug information was available for 32 of the HT women, of whom 21 were taking estrogens alone; 18 of these were taking Premarin, while 3 were taking Estraderm. Eleven women were taking estrogens and a progestin; four were taking Prempro, four were taking Premarin and Provera, two were taking Estrace and Provera, and one was taking

Table 4  
HT use and structure volumes

Structure	Volume (mean $\pm$ S.D.)			HT	
	Male ( <i>n</i> =33)	Non-HT female ( <i>n</i> =51)	HT female ( <i>n</i> =41)	<i>F</i>	<i>P</i>
<b>Gray matter</b>	<b>498.054 <math>\pm</math> 66.432</b>	<b>484.813 <math>\pm</math> 63.010</b>	<b>460.472 <math>\pm</math> 46.967</b>	<b>12.65</b>	<b>&lt;0.0001</b>
<b>White matter</b>	<b>430.250 <math>\pm</math> 59.028</b>	<b>348.784 <math>\pm</math> 57.611</b>	<b>353.513 <math>\pm</math> 45.758</b>	<b>10.60</b>	<b>&lt;0.0001</b>
Ventricular CSF	35.198 $\pm$ 17.036	30.044 $\pm$ 15.020	32.168 $\pm$ 14.049	1.51	0.2241
<b>Nonventricular CSF</b>	<b>266.365 <math>\pm</math> 59.586</b>	<b>221.345 <math>\pm</math> 67.216</b>	<b>248.704 <math>\pm</math> 45.037</b>	<b>7.30</b>	<b>0.0010</b>
White-matter lesions	4.505 $\pm$ 5.814	6.210 $\pm$ 9.387	3.789 $\pm$ 1.997	2.44	0.0912
Left caudate	3.797 $\pm$ 0.726	3.539 $\pm$ 0.582	3.460 $\pm$ 0.551	0.72	0.4875
Right caudate	3.928 $\pm$ 0.756	3.750 $\pm$ 0.628	3.800 $\pm$ 0.478	0.93	0.3948
<b>Left putamen</b>	<b>3.507 <math>\pm</math> 0.626</b>	<b>3.615 <math>\pm</math> 0.710</b>	<b>3.440 <math>\pm</math> 0.511</b>	<b>7.05</b>	<b>0.0013</b>
<b>Right putamen</b>	<b>3.460 <math>\pm</math> 0.770</b>	<b>3.551 <math>\pm</math> 0.657</b>	<b>3.561 <math>\pm</math> 0.664</b>	<b>6.83</b>	<b>0.0015</b>
Left hippocampus	3.167 $\pm$ 0.591	2.862 $\pm$ 0.399	2.832 $\pm$ 0.431	0.39	0.6781
Right hippocampus	3.370 $\pm$ 0.456	3.102 $\pm$ 0.473	2.935 $\pm$ 0.308	1.77	0.1747

The tests of hippocampal volume were conducted with 2 and 68 degrees of freedom. All other tests were conducted with 2 and 120 *df*. All volumes are in cubic centimeters. Values in boldface were significant at  $P < 0.002$ .



Table 5

Pairwise comparisons between men ( $n=33$ ) and non-HT women ( $n=51$ )

Structure	Incremental $R^2$	$t(79)$	$P$	B	S.E.
<b>Gray matter</b>	<b>0.06</b>	<b>3.26</b>	<b>0.0016</b>	<b>38.987</b>	<b>11.969</b>
White matter	0.08	−2.77	0.0068	−37.108	13.408
Nonventricular CSF	<0.01	−0.72	0.4726	−8.414	11.668
<b>Left putamen</b>	<b>0.09</b>	<b>3.58</b>	<b>0.0001</b>	<b>0.450</b>	<b>0.126</b>
<b>Right putamen</b>	<b>0.09</b>	<b>3.23</b>	<b>0.0017</b>	<b>0.458</b>	<b>0.142</b>

All volumes are in cubic centimeters. Values in boldface were significant at  $P<0.002$ .

Ogen and Provera. Dose information was available for 18 women. Sixteen of them were taking Premarin alone; two of these women were taking 0.3 mg/day, while 14 were taking 0.625 mg/day. Two women were taking Prempro at the 0.625/2.5 mg dose. Thus, all of these women were taking conjugated equine estrogens.

We compared the HT group and the non-HT group with one another and with males. Table 2 presents the demographics broken down by group. Analyses of variance revealed no differences in age or education among the three groups;  $t$  tests did not reveal any significant differences in age or education between the HT and non-HT groups (all  $P>0.1$ ).

ICV did not differ between the two groups of women [ $t(90)=-0.37$ ,  $P<0.71$ ] but was larger in men than in either group of women [for men vs. HT,  $t(88)=7.58$ ,  $P<0.0001$ ; for men vs. non-HT,  $t(98)=6.29$ ,  $P<0.0001$ ]. We used Fisher's exact test to check for differences in the frequency of comorbid conditions across groups; we found no significant differences (all  $P>0.05$ ). Table 3 shows the prevalence of comorbid conditions in our participant groups.

We created a set of two effects-coded variables to represent HT status and gender (one level for males, one for non-HT females, and one for HT females), and we entered these into the regression equation after the ICV, age, and education (see Cohen and Cohen, 1988, for a discussion of effects coding in multiple regression) (Table 4).

Table 6

Pairwise comparisons between men ( $n=33$ ) and HT women ( $n=41$ )

Structure	Incremental $R^2$	$t(69)$	$P$	B	S.E.
Gray matter	0.02	2.44	0.0169	31.175	12.792
White matter	0.13	−2.86	0.0053	−43.386	15.175
Nonventricular CSF	0.04	0.33	0.7439	4.073	12.427
Left putamen	0.03	1.86	0.0661	0.247	0.132
Right putamen	0.08	2.60	0.0111	0.464	0.178

All volumes are in cubic centimeters.

Table 7

Pairwise comparisons between HT women ( $n=41$ ) and non-HT women ( $n=51$ )

Structure	Incremental $R^2$	$t(87)$	$P$	B	S.E.
<b>Gray matter</b>	<b>0.07</b>	<b>−3.50</b>	<b>0.0007</b>	<b>−30.277</b>	<b>8.652</b>
White matter	<0.01	−0.27	0.7854	−2.630	9.628
<b>Nonventricular CSF</b>	<b>0.07</b>	<b>3.57</b>	<b>0.0006</b>	<b>31.762</b>	<b>8.887</b>
Left putamen	0.05	−2.33	0.0221	−0.280	0.120
Right putamen	<0.01	−0.63	0.5324	−0.080	0.128

All volumes are in cubic centimeters. Values in boldface were significant at  $P<0.002$ .

The HT variable set was significant for gray matter, white matter, nonventricular CSF, and left and right putamen. We therefore conducted pairwise comparisons on these variables. Tables 5, 6 and 7 present the results of this analysis. In Tables 5 and 6, positive B weights indicate larger volumes in females; in Table 7, positive B weights indicate larger volumes in HT women.

Non-HT women had significantly larger gray-matter volumes than men or HT women. They also had larger left and right putamen volumes than men; men and HT women did not differ from each other on these measures. Non-HT women also had smaller nonventricular CSF volumes than HT women, but they did not differ from men. Men had more white matter than either group of women, although this difference did not quite reach the adjusted significance level. We also compared HT women using estrogen alone with those using estrogens and a progestin, but no significant differences were observed (all  $t<1.86$ ; all  $P>0.07$ ).

We then proceeded to test for differences in neuropsychological scores. As with the brain volumes, we constructed separate models for each test; we first entered age and education followed by the HT variable set. This set was significant only for the Shipley Vocabulary Test [ $F(2,92)=9.18$ ;  $P<0.0002$ ] but not for any other variable (all  $F<4.60$ ; all  $P>0.01$ ). Pairwise analyses revealed that HT females received a higher score than males (for females, mean=35.9, S.D.=3.6; for males, mean=32.8, S.D.=5.9;  $t=3.34$ ,  $P<0.0014$ ). Comparisons of non-HT females and males were nonsignificant, as were comparisons with non-HT females and HT females (for non-HT females, mean=33.8, S.D.=4.3; both  $t<2$ ; both  $P>0.01$ ).

#### 4. Discussion

Several lines of research suggest that HT should have an effect on brain volumes. Estrogen is thought to have a protective effect against dementia (Mortel and Meyer, 1995; Costa et al., 1999; Panidis et al., 2001) and to

reduce the volume (and thus the severity) of a stroke (Hurn and Macrae, 2000), although other studies suggest the opposite conclusion (Shaywitz and Shaywitz, 2000; Bushnell et al., 2001; Women's Health Initiative Writing Group, 2002). In accordance with findings suggesting a protective effect, HT is believed to have beneficial effects on cognitive abilities (see, for example, Steffens et al., 1999; Miller et al., 2002), particularly various forms of memory (Robinson et al., 1994; Sherwin, 1999; Maki et al., 2001), although numerous studies have failed to find any benefit of HT on cognition (Barrett-Connor and Kritz-Silverstein, 1993; de Moraes et al., 2001). The exact mechanism behind estrogen's protective effects remains unclear, but it is believed to have an effect on the hippocampus (Azcoitia et al., 1999) as well as the prefrontal cortex (Keenan et al., 2001), perhaps by increasing glucose metabolism (Rasgon et al., 2001) and reducing the sensitivity of the hypothalamic–pituitary–adrenal axis (Seeman et al., 2001). So far, however, little is known about the effects of HT on brain volumes. One study found no differences in white-matter lesions but observed that women on HT had slightly more brain atrophy than non-HT women (Luoto et al., 2000); another found that women on HT had larger hippocampal volumes (Eberling et al., 2003).

In general, our analysis showed that the brains of HT women showed substantially more severe atrophy than those of non-HT women. On two measures (gray matter volume and nonventricular CSF volume), the adjusted brain volumes of women on HT were not significantly different from those of men, but were significantly different from those of non-HT women. The gray-matter volumes of HT women were over 30 cm<sup>3</sup> smaller than those of their non-HT counterparts, a difference that amounts to over 6% of the mean female volume. Non-ventricular CSF was larger in HT women by 32 cm<sup>3</sup>, or 14% of the mean female volume. Other comparisons did not follow this trend, showing differences between one group of women and men, but not between the two groups of women. White-matter volumes were larger in HT women than in men by over 40 cm<sup>3</sup> (about 10% of the overall average volume); the two groups of women did not differ on this measure. Men and non-HT women differed on several measures: men had less gray matter, a smaller left and right putamen, and more white matter. We found no differences in any other measures, including the hippocampus; however, fewer hippocampal volumes were available for our analysis, which may have prevented us from seeing a significant result.

Overall, our examination of hormone therapy produced unanticipated results. Much of the literature suggests that the brains of women on HT should show

less atrophy than those of nonusers. Nevertheless, when our analyses found any differences at all, they tended to find the reverse: HT women had less gray matter and more nonventricular CSF than non-HT women. These differences cannot be attributed to differences in education or age, since these factors were nearly identical across groups.

Previous studies have identified the health of the participants as a potential confounding factor: one hypothesis suggests that HT women start taking HT because they are initially less healthy, but another common hypothesis, is however, which states that HT users may perform better on many health measures simply because they are more attentive to their health than nonusers (e. g. Matthews et al., 1996). Our analyses indicated that each of our groups was equally healthy, suggesting that medical history cannot account for the difference we observed among groups.

Our results differed from those of previous studies (Luoto et al., 2000; Eberling et al., 2003). Some of these differences could stem from differences in the participant populations. Participants in our HT group had been taking HT for at least 10 years, and those in our non-HT group had never taken HT. Eberling et al. (2003) selected participants based on their current use of HT and did not investigate medical histories. If HT use is in fact harmful, perhaps longer use is associated with poorer outcomes. We did not study this factor directly, though, and therefore cannot come to any firm conclusions; moreover, Luoto et al. (2000) found no effect of duration of HT. The type of HT also differed across these studies; a greater proportion of our participants was taking a combination of estrogens and a progestin. The relevance of this difference is uncertain, however. Both kinds of therapy are thought to increase stroke risk (Women's Health Initiative Writing Group, 2002; Women's Health Initiative Steering Committee, 2004). In any event, the two previous studies had similar proportions of combination-therapy users—but they nevertheless yielded conflicting results.

Our findings should be interpreted with some caution, because we were unable to examine women before and after HT treatment, so we do not know if HT women simply started with smaller volumes. Our small sample size precluded a thorough examination of the effects of different drugs, drug combinations, doses, and routes of administration; we were also unable to collect sufficient information about hysterectomies, date and duration of menopause, or childbearing history. All of these factors can affect outcomes and may have accounted in part for the unusual results we observed. Second, the importance of these findings is unclear, given that the differences in

brain volumes do not seem to be reflected in neuropsychological test scores. A definitive study would involve a standard double-blind placebo-controlled design that examined brain volumes longitudinally, starting before the onset of therapy and ending after several years of HT. Given the recent controversy over the efficacy and safety of HT (Women's Health Initiative Writing Group, 2002; Women's Health Initiative Steering Committee, 2004), however, this design may not be a viable option.

## Acknowledgments

The authors thank Mr. Timothy Blitchington for the development of the GRID Program, Dr. Christopher Byrum for assistance in developing neuroanatomical guidelines, and Dr. Kathleen A. Welsh-Bohmer for developing the neuropsychological test battery. This study was supported by NIMH grants P50 MH60451, P30 MH40159, R01 MH54846, and K07 MH01367.

## References

- Azcoitia, I., Sierra, A., Garcia-Segura, L.M., 1999. Neuroprotective effects of estradiol in the adult rat hippocampus: interaction with insulin-like growth factor-I signaling. *Journal of Neuroscience Research* 58, 815–822.
- Barrett-Connor, E., Kritz-Silverstein, D., 1993. Estrogen replacement therapy and cognitive function in older women. *Journal of the American Medical Association* 269, 2637–2641.
- Benton, A.L., 1974. Revised Visual Retention Test, 4th ed. Psychological Corporation, New York.
- Benton, A.L., Hamsher, K., Varney, N., Spreen, O., 1983. Contributions to Neuropsychological Assessment. Oxford University Press, New York.
- Breitner, J.C.S., Welsh, K.A., Robinette, C.D., 1994. Alzheimer's disease in the National Academy of Sciences Registry of aging twin veterans: II. Longitudinal findings in a pilot series. *Dementia* 5, 99–105.
- Bushnell, C.D., Samsa, G.P., Goldstein, L.B., 2001. Hormone replacement therapy and ischemic stroke severity in women: a case-control study. *Neurology* 56, 1304–1307.
- Byrum, C.E., MacFall, J.R., Charles, H.C., Chitilla, V.R., Boyko, O.B., Upchurch, L., Smith, J.S., Ragahopalan, P., Passe, T., Kim, D., Xanthakos, S., Krishnan, K.R.R., 1996. Accuracy and reproducibility of brain and tissue volumes using a magnetic resonance segmentation method. *Psychiatry Research: Neuroimaging* 67, 215–234.
- Clark, C.M., Ewbank, D.C., 1996. Performance of the Dementia Severity Rating Scale: a caregiver questionnaire for rating severity in Alzheimer disease. *Alzheimer's Disease and Associated Disorders* 10, 31–39.
- Cohen, J., Cohen, P., 1988. Applied multiple regression/correlation analysis for the behavioral sciences, 2nd ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Cook, I.A., Morgan, M.L., Dunkin, J.J., David, S., Witte, E., Lufkin, R., Abrams, M., Rosenberg, S., Leuchter, A.F., 2002. Estrogen replacement therapy is associated with less progression of subclinical structural brain disease in normal elderly women: a pilot study. *International Journal of Geriatric Psychiatry* 17, 610–618.
- Costa, M.M., Reus, V.I., Wolkowitz, O.M., Manfredi, F., Lieberman, M., 1999. Estrogen replacement therapy and cognitive decline in memory-impaired post-menopausal women. *Biological Psychiatry* 46, 182–188.
- Cowell, P.E., Turetsky, B.I., Gur, R.C., Grossman, R.I., Shtasel, D.L., Gur, R.E., 1994. Sex differences in aging of the human frontal and temporal lobes. *Journal of Neuroscience* 14, 4748–4755.
- De Moraes, S.A., Szklo, M., Knopman, D., Park, E., 2001. Prospective assessment of estrogen replacement therapy and cognitive functioning—atherosclerosis risk in communities study. *American Journal of Epidemiology* 154, 733–739.
- Eberling, J.L., Wu, C., Haan, M.N., Mungas, D., Buonocore, M., Jagust, W.J., 2003. Preliminary evidence that estrogen protects against age-related hippocampal atrophy. *Neurobiology of Aging* 24, 725–732.
- Golomb, J., de Leon, M.J., Kluger, A., George, A.E., Tarshish, C., Ferris, S.H., 1993. Hippocampal atrophy in normal aging. An association with recent memory impairment. *Archives of Neurology* 50, 967–973.
- Greenberg, D.L., Fetzer, D.L., Payne, M.E., MacFall, J.R., Provenzale, J.M., Steffens, D.C., Krishnan, R.R., submitted for publication. Aging, gender, and the elderly adult brain: a comparison of analytical strategies.
- Gunning-Dixon, F.M., Head, D., McQuain, J., Acker, J.D., Raz, N., 1998. Differential aging of the human striatum: a prospective MR imaging study. *American Journal of Neuroradiology* 19, 1501–1507.
- Gur, R.C., Mozley, P.D., Resnick, S.M., Gottlieb, G.L., Kohn, M., Zimmerman, R., Herman, G., Atlas, S., Grossman, R., Beretta, D., 1991. Gender differences in age effect on brain atrophy measured by magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America* 88, 2845–2849.
- Gur, R.C., Gunning-Dixon, F.M., Turetsky, B.I., Bilker, W.B., Gur, R.E., 2002. Brain region and sex differences in age association with brain volume: a quantitative MRI study of healthy young adults. *American Journal of Geriatric Psychiatry* 10, 72–80.
- Hoppe, C.D., Muller, U.D., Werheid, K.D., 2000. Digit Ordering Test: clinical, psychometric, and experimental evaluation of a verbal working memory test. *Clinical Neuropsychology* 14, 38–55.
- Hurn, P.D., Macrae, I.M., 2000. Estrogen as a neuroprotectant in stroke. *Journal of Cerebral Blood Flow and Metabolism* 20, 499–504.
- Kaplan, E.F., Goodglass, H., Weintraub, S., 1983. The Boston Naming Test, 2nd ed. Lea & Febiger, Philadelphia, Pa.
- Keenan, P.A., Ezzat, W.H., Ginsburg, K., Moore, G.J., 2001. Prefrontal cortex as the site of estrogen's effect on cognition. *Psychoneuroendocrinology* 26, 577–590.
- Kikinis, R., Shenton, M., Jolesz, F., Gerig, G., Martin, J., Anderson, M., Metcalf, D., Guttmann, C., McCarley, R., Lorensen, W., Cline, H., 1992. Quantitative analysis of brain and cerebrospinal fluid spaces with MR imaging. *Journal of Magnetic Resonance Imaging* 2, 619–629.
- Luoto, R., Manolio, T., Meilahn, E., Bhadelia, R., Furberg, C., Cooper, L., Kraut, M., 2000. Estrogen replacement therapy and MRI-demonstrated cerebral infarcts, white matter changes, and brain atrophy in older women: the Cardiovascular Health Study. *Journal of the American Geriatric Society* 48, 467–472.
- MacFall, J.R., Byrum, C.E., Parashos, I., Early, B., Charles, H.C., Chitilla, V., Boyko, O.B., Upchurch, L., Krishnan, K.R., 1994. Relative accuracy and reproducibility of regional MRI brain

- volumes for point-counting methods. *Psychiatry Research* 55, 167–177.
- Maki, P., Zonderman, A., Resnick, S., 2001. Enhanced verbal memory in nondemented elderly women receiving hormone-replacement therapy. *American Journal of Psychiatry* 158, 227–233.
- Mathews, K.A., Kuller, L.H., Wing, R.R., Meilahn, E.N., Plantinga, P., 1996. Prior to use of estrogen replacement therapy, are users healthier than nonusers? *American Journal of Epidemiology* 143, 971–978.
- Miller, K.J., Conney, J.C., Rasgon, N.L., Fairbanks, L.A., Small, G.W., 2002. Mood symptoms and cognitive performance in women estrogen users and nonusers and men. *Journal of the American Geriatric Society* 50, 1826–1830.
- Morris, J.C., Heyman, A., Mohs, R.C., 1989. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I: clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 39, 1159–1165.
- Mortel, K.F., Meyer, J.S., 1995. Lack of postmenopausal estrogen replacement therapy and the risk of dementia. *Journal of Neuropsychiatry and Clinical Neurosciences* 7, 334–337.
- Murphy, D.G., DeCarli, C., McIntosh, A.R., Daly, E., Mentis, M.J., Pietrini, P., Szczepanik, J., Schapiro, M.B., Grady, C.L., Horwitz, B., Rapaport, S.I., 1996. Sex differences in human brain morphometry and metabolism: an in vivo quantitative magnetic resonance imaging and positron emission tomography study on the effect of aging. *Archives of General Psychiatry* 53, 585–594.
- Paganini-Hill, A., 2001. Hormone replacement therapy and stroke: risk protection or no effect? *Maturitas* 38, 243–261.
- Panidis, D.K., Matalliotakis, I.M., Rousso, D.H., Kourtis, A.I., Koumantakis, E.E., 2001. The role of estrogen replacement therapy in Alzheimer's disease. *European Journal of Obstetrics and Gynecology* 95, 86–91.
- Payne, M.E., Fetzter, D.L., MacFall, J.R., Provenzale, J.M., Byrum, C.E., Krishnan, K.R., 2002. Development of a semi-automated method for quantification of MRI gray and white matter lesions in geriatric subjects. *Psychiatry Research: Neuroimaging* 115, 63–77.
- Rasgon, N.L., Small, G.W., Siddarth, P., Miller, K., Ercoli, L.M., Bookheimer, S.Y., Lavretsky, H., Huang, S.C., Barrio, J.R., Phelps, M.E., 2001. Estrogen use and brain metabolic change in older adults. A preliminary report. *Psychiatry Research: Neuroimaging* 107, 11–18.
- Raz, N., Torres, I.J., Acker, J.D., 1995. Age, gender, and hemispheric differences in human striatum: a quantitative review and new data from in vivo MRI morphometry. *Neurobiology of Learning and Memory* 63, 133–142.
- Raz, N., Gunning, F.M., Head, D., Dupuis, J.H., McQuain, J., Briggs, S.D., Loken, W.J., Thornton, A.E., Acker, J.D., 1997. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cerebral Cortex* 7, 268–282.
- Reitan, R.M., 1992. Trail Making Test: Manual for Administration and Scoring. Reitan Neuropsychological Laboratory, Tucson, AZ.
- Resnick, S.M., Goldszal, A.F., Davatzikos, C., Golski, S., Kraut, M.A., Metter, E.J., Bryan, R.N., Zonderman, A.B., 2000. One-year age changes in MRI brain volumes in older adults. *Cerebral Cortex* 10, 464–472.
- Resnick, S.M., Pham, D.L., Kraut, M.A., Zonderman, A.B., Davatzikos, C., 2003. Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *Journal of Neuroscience* 23, 3295–3301.
- Robinson, D., Friedman, L., Marcus, R., Tinklenberg, J., Yesavage, J., 1994. Estrogen replacement therapy and memory in older women. *Journal of the American Geriatric Society* 42, 919–922.
- Robinson-Whelen, S., Storandt, M., 1992. Immediate and delayed prose recall among normal and demented adults. *Archives of Neurology* 49, 32–34.
- Schmidt, R., Fazekas, F., Reinhart, B., Kappeler, P., Fazekas, G., Offenbacher, H., Eber, B., Schumacher, M., Freidl, W., 1996. Estrogen replacement therapy in older women: a neuropsychological and brain MRI study. *Journal of the American Geriatrics Society* 44, 1307–1313.
- Seeman, T.E., Singer, B., Wilkinson, C.W., McEwen, B., 2001. Gender differences in age-related changes in HPA axis reactivity. *Psychoneuroendocrinology* 26, 225–240.
- Shaywitz, B.A., Shaywitz, S.E., 2000. Estrogen and Alzheimer disease—plausible theory, negative clinical trial. *Journal of the American Medical Association* 283, 1055–1056.
- Sherwin, B.B., 1999. Can estrogen keep you smart? Evidence from clinical studies. *Journal of Psychiatry and Neuroscience* 24, 315–321.
- Smith, A., 1982. Symbol Digit Modalities Test-Manual. Western Psychological Services, Los Angeles, CA.
- Steffens, D.C., Norton, M.C., Plassman, B.L., Tschanz, J.T., Wyse, B.W., Welsh-Bohmer, K.A., 1999. Enhanced cognitive performance with estrogen use in nondemented community-dwelling older women. *Journal of the American Geriatric Society* 47, 1171–1175.
- Steffens, D.C., Byrum, C.E., McQuoid, D.R., Greenberg, D.L., Payne, M.E., Blitchington, T.F., MacFall, J.R., Krishnan, K.R., 2000. Hippocampal volume in geriatric depression. *Biological Psychiatry* 48, 301–309.
- Wechsler, D., 1981. The Wechsler Adult Intelligence Scale Revised. Psychological Corporation, New York.
- Wechsler, D., 1987. Wechsler Memory Scale-Revised Manual. Text: Psychological Corporation, San Antonio.
- Welsh, K.A., Butters, N., Mohs, R.C., 1994. The Consortium to Establish a Registry of Alzheimer's Disease (CERAD). Part V: a normative study of the neuropsychological battery. *Neurology* 44, 609–614.
- Women's Health Initiative Steering Committee, 2004. The effects of conjugated equine estrogen in postmenopausal women with hysterectomy. *Journal of the American Medical Association* 291, 1701–1712.
- Women's Health Initiative Writing Group, 2002. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. *Journal of the American Medical Association* 288, 321–333.
- Xu, J., Kobayashi, S., Yamaguchi, S., Iijima, K., Okada, K., Yamashita, K., 2000. Gender effects on age-related changes in brain structure. *American Journal of Neuroradiology* 21, 112–118.
- Zachary, R.A., 1991. Shipley Institute of Living Scale-Revised Manual. Western Psychological Services, Los Angeles, CA.
- Zandi, P.P., Carlson, M.C., Plassman, B.L., Welsh-Bohmer, K.A., Mayer, L.S., Steffens, D.C., Breitner, J.C., 2002. Hormone replacement therapy and incidence of Alzheimer disease in older women: the Cache County study. *Journal of the American Medical Association* 288, 2123–2129.