Predicting memory decline in normal elderly: Genetics, MRI, and cognitive reserve

Larry A. Tupler a,∗, K. Ranga R. Krishnan a, Daniel L. Greenberg a,1, Santica M. Marcovina b, Martha E. Payne a, James R. MacFall c, H. Cecil Charles c, P. Murali Doraiswamy a

a Department of Psychiatry and Behavioral Sciences, Box 3018, Duke University Medical Center, Durham, NC 27710, USA
b Northwest Lipid Metabolism and Diabetes Research Laboratories, University of Washington, Seattle, WA 98109, USA
c Department of Radiology, Duke University Medical Center, Durham, NC 27710, USA

Received 4 January 2006; received in revised form 23 June 2006; accepted 6 July 2006

Available online 17 August 2006

Abstract

Major predictors of Alzheimer’s disease (AD) include apolipoprotein E (APOE)-ε4, hippocampal atrophy on magnetic resonance imaging (MRI), and memory dysfunction prior to diagnosis. We examined 159 normal elderly subjects with MRI and the California Verbal Learning Test (CVLT); 84 returned for longitudinal follow-up 5 years later. Analyses at baseline revealed significant variance in hippocampal volume accounted for by cerebral volume and age but not by APOE isoform. However, interactions involving APOE isoform and laterality were observed. As hypothesized, an APOE × time interaction was revealed for CVLT long-delay free recall: APOE-ε3/4 subjects had significantly poorer performance than APOE-ε3/3 subjects at follow-up. Forward stepwise multiple regression analysis predicting follow-up long-delay free recall selected baseline recall, followed by number of APOE-ε4 alleles, followed by left-hippocampal volume. Age and sex did not enter into the model. We conclude that APOE-ε4 predicts longitudinal memory decline in healthy controls and that MRI morphometry of hippocampus adds slightly to predictive value.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Alzheimer’s disease; Apolipoprotein E; California Verbal Learning Test; Cerebral volume; Hippocampus; Memory; Morphometry; MRI; Normal aging

1. Introduction

Major predictors of Alzheimer’s disease (AD) include apolipoprotein E (APOE)-ε4 [8–10,39,40,52], hippocampal atrophy [1,16,33,41–43,50,51,59,76,102], and memory dysfunction prior to diagnosis. We examined 159 normal elderly subjects with MRI and the California Verbal Learning Test (CVLT); 84 returned for longitudinal follow-up 5 years later. Analyses at baseline revealed significant variance in hippocampal volume accounted for by cerebral volume and age but not by APOE isoform. However, interactions involving APOE isoform and laterality were observed. As hypothesized, an APOE × time interaction was revealed for CVLT long-delay free recall: APOE-ε3/4 subjects had significantly poorer performance than APOE-ε3/3 subjects at follow-up. Forward stepwise multiple regression analysis predicting follow-up long-delay free recall selected baseline recall, followed by number of APOE-ε4 alleles, followed by left-hippocampal volume. Age and sex did not enter into the model. We conclude that APOE-ε4 predicts longitudinal memory decline in healthy controls and that MRI morphometry of hippocampus adds slightly to predictive value.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Alzheimer’s disease; Apolipoprotein E; California Verbal Learning Test; Cerebral volume; Hippocampus; Memory; Morphometry; MRI; Normal aging
but not all [4, 7, 14, 15, 25, 45, 61, 67, 72, 73, 80] cross-sectional studies of non-demented middle-aged and elderly subjects with APOE-ε4 have reported poorer learning and memory [8, 9, 13, 21, 27, 44, 56, 78, 84, 92] and other impairments in cognitive performance [6, 32, 44, 82, 101].

Longitudinal MRI studies of hippocampal changes in association with APOE-ε4 have been relatively sparse but have suggested that APOE-ε4 is associated with accelerated hippocampal-volume loss in healthy controls [15, 63]. Most neuropsychological studies have shown that APOE-ε4 is associated with declining memory [2, 6, 23, 36, 38, 44, 61, 88] and other cognitive processes [6, 20, 23, 44, 88], but null findings have also been reported [11, 15, 52, 72]. Some studies have reported no cross-sectional differences in memory but significant longitudinal differences [6, 61, 67, 79, 89]; others have reported the reverse [101]. One study reported a decline in cognition for women only [64]. Dik et al. [22] reported that APOE-ε4 was associated with memory decline in cognitively impaired subjects but not in normal subjects after a 3-year interval. Findings relating APOE-ε4 to memory performance in healthy individuals are therefore equivocal.

Although a number of studies have separately examined APOE-ε4 in relation to hippocampal volume or to memory, few studies have evaluated the predictive value of all three in combination. Marquis et al. [59] found hippocampal volume and memory performance to predict the development of questionable dementia, but APOE-ε4 did not add to the predictive value. The relative contribution of hippocampal volume versus memory performance to diagnosis is also uncertain, with some investigators questioning the value of hippocampal morphometry [50, 85]. The use of APOE genotyping alone in the diagnosis of AD is not supported [60, 94, 97], and thus combination studies are warranted examining the relative value of APOE isoform status and hippocampal morphometry in predicting memory decline.

We therefore studied the relationship between APOE-ε4, hippocampal volume, and both cross-sectional and longitudinal memory performance over a 5-year interval. Because delayed recall has been shown to be the most sensitive memory-related predictor of the development of AD [8, 47, 48, 56, 100], we hypothesized that possession of APOE-ε4 would be related to a longitudinal decline in California Verbal Learning Test (CVLT) long-delay free recall [8, 9]. We also hypothesized that hippocampal volume at baseline would predict CVLT performance at follow-up and that the combination of APOE-ε4 and hippocampal volume would be superior in predicting follow-up memory scores relative to APOE-ε4 alone.

2. Materials and methods

2.1. Subjects

The study was approved by the Institutional Review Board of Duke University Medical Center, and each participant provided written informed consent following description of the procedures. Subjects consisted of 163 men and women ages 55–85, of which 153 were described in Chen et al. [14]. All subjects were evaluated by a physician and study coordinator. This evaluation included a neuropsychiatric examination and a review of medical records to qualify subjects as normal community-dwelling volunteers. APOE genotyping was performed with DNA extracted from buffy coat, using polymerase chain reaction (PCR) amplification and the HhaI restriction-digest method of Hixson and Vernier [37]. Group classifications consisted of subjects with an APOE-ε4 load of 0 (APOE-ε2/3 or APOE-ε3/3), 1 (APOE-ε2/4 or APOE-ε3/4), or 2 (APOE-ε4/4) ε4 alleles.

Demographic, mental-status, and AD family-history characteristics of the sample as a function of APOE allele combination are presented in Table 1. No significant age differences were observed between the groups (F[4,158] = 0.61, p > 0.65). Because of the low number of subjects in the APOE-ε2/4 group (n = 4), they were excluded from further analysis, leaving a final sample of 159. Subjects were predominantly Caucasian. No differences were observed between APOE groups in sex composition (Fisher’s exact p > 0.61), education level (F[3,154] = 1.13, p > 0.33), or Mini-Mental State Examination (MMSE) score (F[3,154] = 1.80, p > 0.14). Borderline significance was observed for endorsement of a family history of AD (Fisher’s exact p = 0.09). One hundred and fifty-seven subjects received MRI, of which 150 had hippocampal measurements; 157 had CVLT data.

Follow-up subjects consisted of 84 of these individuals tested approximately 5 years later. As with the baseline sample, no significant differences were observed for APOE groups with respect to age (F[3,80] = 0.18, p > 0.90) or education (F[3,80] = 0.21, p > 0.89) at study entry. Mean Beck Depression Inventory [55] scores at follow-up were in the normal range (APOE-ε2/3 = 6.5 ± 6.0, ε3/3 = 5.4 ± 3.7, ε3/4 = 6.0 ± 4.0, ε4/4 = 4.2 ± 2.4) and did not differ between groups (F[3,79] = 0.51, p > 0.67). However, the groups did differ in their sex composition (Fisher’s exact p < 0.009), with relatively more females:males in the APOE-ε2/3 (7:1) and -ε3/3 (25:14) groups and fewer in the APOE-ε3/4 (12:19) and -ε4/4 (1:5) groups. Eighty-two subjects had MRI data; all 84 subjects received the CVLT.

No differences in frequencies of APOE allele combinations were observed between those who were studied longitudinally (8 ε2/3, 39 ε3/3, 31 ε3/4, 6 ε4/4) and those lost to follow-up (11 ε2/3, 32 ε3/3, 29 ε3/4, 3 ε4/4; Fisher’s exact p > 0.64). Likewise, sex compositions were statistically equivalent for those who were studied (45 F, 39 M) and those lost to follow-up (38 F, 37 M; χ²[1] = 0.13, p > 0.71). No differences were observed in the mean age of those studied (65.3 ± 6.3 years) and those lost to follow-up (66.3 ± 7.9 years; t[141] = 0.90, p > 0.37) or in the level of education for those studied (15.6 ± 2.6 years) and those lost to follow-up (15.6 ± 2.9 years; t[156] = 0.03, p > 0.97). MMSE score was equivalent for those studied (28.2 ± 1.7) compared with those lost to follow-up (28.4 ± 1.4; t[156] = 0.45, p > 0.65),
Table 1
Baseline sample characteristics of normal community-dwelling volunteers for age, sex, race, level of education, MMSE score, and family history of AD (n = 163)

<table>
<thead>
<tr>
<th>APOE</th>
<th>e2/3</th>
<th>e3/3</th>
<th>e2/4</th>
<th>e3/4</th>
<th>e4/4</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>19</td>
<td>71</td>
<td>4</td>
<td>60</td>
<td>9</td>
</tr>
<tr>
<td>Age (years)</td>
<td>66.9 ± 7.8</td>
<td>65.9 ± 6.7</td>
<td>65.9 ± 6.4</td>
<td>64.9 ± 7.2</td>
<td>68.2 ± 8.0</td>
</tr>
<tr>
<td>Females (n)</td>
<td>11</td>
<td>39</td>
<td>2</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>Males (n)</td>
<td>8</td>
<td>32</td>
<td>2</td>
<td>30</td>
<td>6</td>
</tr>
<tr>
<td>Race(Caucasian/African–American/Hispanic)</td>
<td>17/2/0</td>
<td>68/1/1</td>
<td>4/0/0</td>
<td>57/3/0</td>
<td>9/0/0</td>
</tr>
<tr>
<td>Education (years)</td>
<td>15.9 ± 3.0</td>
<td>15.6 ± 2.7</td>
<td>17.8 ± 4.0</td>
<td>15.2 ± 2.6</td>
<td>16.9 ± 3.9</td>
</tr>
<tr>
<td>MMSE score (total)</td>
<td>28.8 ± 0.9</td>
<td>28.3 ± 1.8</td>
<td>29.3 ± 1.0</td>
<td>28.3 ± 1.4</td>
<td>27.3 ± 2.0</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>7</td>
<td>10</td>
<td>1</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>6</td>
<td>21</td>
<td>1</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>5</td>
<td>21</td>
<td>2</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Family history of AD (% positive)</td>
<td>6 (32%)</td>
<td>27 (38%)</td>
<td>2 (50%)</td>
<td>33 (55%)</td>
<td>6 (67%)</td>
</tr>
</tbody>
</table>

and the number of subjects with/without a family history of AD was equivalent for those followed (39/45) compared with those lost to follow-up (33/41; $\chi^2_{[1]} = 0.05$, p > 0.81). Baseline CVLT performance, however, differed on 1 out of 10 measures, list B, which was slightly higher in those studied (6.4 ± 2.1 words) compared with those lost to follow-up (5.7 ± 1.9 words; $t_{[155]} = 2.05$, p < 0.05).

2.2. MRI acquisition and analysis

Subjects were scanned 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 canthomeatal line and the sagittal lights were aligned with the center of the nose. A rapid sagittal localizer scan was acquired to confirm the alignment.

2.2.1. High-resolution imaging for volume measurements

Two acquisitions were obtained in the axial plane: a three-dimensional (3D) inversion-recovery-prepped spoiled gradient (SPGR)-echo scan for measurement of the hippocampus and a dual-echo fast spin-echo (FSE) acquisition for morphometry of all other cerebral structures. Pulse-sequence parameters for the FSE scan were TR = 4000 ms; TE = 20, 100 ms; nominal bandwidth ±16 kHz; echo-train length = 8 per echo; 256 × 256 matrix; 2.5-mm section thickness; 1 Nex; 24-cm FOV. The images were acquired in two separate acquisitions with a 2.5-mm gap between sections for each acquisition. The second acquisition was offset by 2.5 mm from the first so that the resulting dataset consisted of contiguous sections. For the 3D-SPGR sequence, the parameters were TR = 11.3 ms, TE = 2.2 ms, TI = 300 ms, nominal bandwidth ±16 kHz, flip 20°, 256 × 256 × 124 matrix, 1.5-mm section thickness, 1 Nex, and 24-cm FOV.

2.2.2. MRI processing for brain volumes

The MRI data were transferred to the Neuropsychiatric Imaging Research Laboratory (NIRL), located at Duke University Medical Center, for processing on SUN workstations. Thin-slice, dual-echo FSE images, consisting of proton-density and T2-weighted images, were used for all processing except the hippocampus. Hippocampal measurements utilized the 3D-SPGR sequence. Two computer programs were used to perform volume measurements. Cerebral-volume measurements used a NIRL-modified version of MrX software, which was created by GE Corporate Research and Development (Schenectady, NY) and originally modified for image segmentation by Brigham and Women’s Hospital (Boston, MA). Hippocampal volumes were acquired using the GRID program, which was developed at NIRL. All measurements were performed blind to APOE and memory-performance data.

Cerebral volume was obtained with a segmentation protocol used by NIRL that was a modified version of that developed by Kikinis et al. [46] and Byrum et al. [12] as described previously [70]. This method was a supervised, semi-automated procedure that used the multiple MRI contrasts available to identify different tissue classifications through a ‘seeding’ process wherein a trained analyst manually selected pixels in each tissue type that was to be identified (such as gray matter, white matter, cerebrospinal fluid [CSF], background). The process then assigned a unique value (an integer from 1 to 9) to the identified tissue to produce a new image (‘segmented’ image). Some modifications to Byrum et al.’s [12] method were performed for this study to optimize segmentation of geriatric brain scans. The seed-
2.2.3. Hippocampal measurements

Hippocampal measurements were based on established guidelines [24,75] and were performed on 3D scans to increase the ability to detect and delineate the boundaries of the hippocampus [75]. The anterior commissure (AC) and posterior commissure (PC) were identified in the sagittal view, and a line was drawn between them (the AC–PC line); the scan was then realigned to this plane. The realignment was confirmed by inspecting the axial plane to ensure that the AC and the PC appeared on the same slice.

On each scan, the hippocampus was traced beginning with the most superior axial slice and then proceeding inferiorty. The first slice was defined as the most superior slice on which the colliculi and cerebral aqueduct both appeared. For the last slice, the coronal view was used to locate the slice that was tangent to the most inferior portion of the hippocampus. This slice was excluded, and the slices above it were measured. If one slice was obscured by motion or poor contrast, a volume for that slice was interpolated by averaging the volumes of the previous and subsequent slices. If more than one slice was obscured, the scan was rejected. Blood vessels were transected if they extended into the body of the hippocampus but excluded if they followed along the body. Hyperintense regions within the hippocampus were included; gray-matter regions within the CSF of the temporal horn were excluded. If it was unclear whether a particular point was part of the hippocampus, that point was examined on the coronal and sagittal scans; the point was included if it was in the hippocampus in one or both views.

Superior region (0–3 slices). On these slices, the hippocampus appears as a small, kidney-shaped structure on the medial side of the posterior lateral ventricles. Tracing began at the posterior lateral corner of the hippocampus. It then proceeded medially along the gray/white boundary that forms the posterior border of the hippocampus. If the parahippocampal gyrus was visible between the white matter and the cistern, tracing proceeded to the most anterior white-matter point of the tissue surrounding the cistern; a horizontal line was drawn from this point to the cistern. If the parahippocampal gyrus was not visible, then tracing proceeded directly to the cistern. From this point, tracing continued anteriorly along the medial border of the hippocampus between the CSF and the hippocampal gray matter. A horizontal line was drawn from the CSF of the cistern to the most anterior point of the lateral ventricle. Tracing then proceeded along the CSF/gray-matter or white-matter/gray-matter border and returned to the starting point.

Middle region (0–2 slices). This step was used if the gray matter of the hippocampus extended past the posterior ventricles but the amygdalo-hippocampal transition zone was not yet visible. Tracing began at the posterior lateral corner of the hippocampus and continued to the cistern as described in the previous step. It then proceeded along the CSF/gray-matter border; however, if the hippocampus was inseparable from the pontine white matter, tracing proceeded along the hippocampal-pontine border. The anterior border was delineated by tracing along the posterior border of the amygdala out to the white matter. Tracing then proceeded along the CSF/gray-matter or white-matter/gray-matter border and returned to the starting point.

Transition zone (1–3 slices). This step was used when the temporal horn of the lateral ventricle was visible and the amygdalo-hippocampal transition zone appeared as a strip of gray matter connecting the hippocampus to the amygdala. Tracing began at the posterior lateral border of the hippocampus, then proceeded to the cistern and along the medial border as in the previous steps. The transition zone was transected by drawing a horizontal line from the cistern (or the pontine white matter) to the most medial point of the temporal horn of the lateral ventricle. Tracing then continued along the posterior border of the amygdala out to the white matter, proceeded along the lateral border, and returned to the starting point.

Inferior region (1–3 slices). This step was used when the transition zone was no longer visible. Tracing began at the posterior lateral border of the hippocampus, then continued along the boundary between the gray and white matter that formed the posterior border of the hippocampus. If a strip of white matter separated the hippocampal body from the parahippocampal gyrus, tracing continued around the gray-matter/white-matter border. If, on the other hand, the hippocampal body was connected to the parahippocampal gyrus, then tracing proceeded to the cistern and along the medial border as in previous steps. In either case, tracing continued...
along the gray-matter/white-matter border and returned to the starting point.

2.2.4. Training and reliability
All NfMRI image-analysis technicians received extensive training by experienced analysts. Reliability was established by repeated measurements on multiple MRI scans before raters were approved to process study data. Inter-rater intraclass correlation coefficients (ICCs) attained were as follows: total brain = 0.998, left cerebral hemisphere = 0.996, right cerebral hemisphere = 0.997, total cerebrum = 0.997. All hippocampal measurements were performed by one of the authors (DLG). Intra-rater ICCs were 0.88 for left hippocampus and 0.84 for right hippocampus.

2.3. CVLT administration
Subjects were administered the CVLT [7,9,14,48,52,69] under standard administration conditions in a quiet examination room. The interval between short-delay and long-delay recall was interpolated with visuospatial and attentional tasks such as Wechsler Adult Intelligence Scale-Revised (WAIS-R) block design, WAIS-R digit symbol substitution, Ruff Figural Fluency Test, 2 & 7 Test of Visual Attention, Trail Making Test, and/or the Benton Facial Recognition Test [57].

2.4. Statistical analyses
Statistics were computed using SAS Version 8.2 (SAS Institute Inc., Cary, NC). Cross-sectional analyses of baseline measures examined differences in total cerebral volume between APOE groups (ε2/3, ε3/3, ε3/4, ε4/4) using a general linear models (GLM) type III analysis, with APOE group as a class variable, sex coded as a dummy variable, and computation of an APOE × sex interaction term. Hippocampal-volume measurements were then examined with a GLM analysis using APOE group as a class variable and cerebral volume, age, sex (dummy coded), and the APOE × sex interaction as control variables examined for their contribution to the variance. Volumetric results are reported in both raw form and as least-square means adjusted for the covariate terms in the model. Relationships between volume measures and CVLT performance were then examined using Spearman rank-order correlations and GLM analyses examining total volume as recalled as the dependent measure regressed against APOE-ε4 load (i.e., number of APOE-ε4 alleles), sex, MRI volume, and the APOE-ε4 load × volume interaction.

Longitudinal changes in CVLT scores over time were examined in a repeated-measures analysis of covariance (ANCOVA), with APOE group as a class variable, sex as a control variable, and time (baseline and follow-up) as a repeated measure. Significant differences between means were subsequently examined using post hoc Tukey pairwise comparisons. We were primarily interested in measures of long-delay free recall. We secondarily examined learning by analyzing performance on trial 1 and trial 5 in a repeated-measures ANCOVA, with the effects of trial and time both serving as repeated measures. Secondary analyses also examined total words recalled for trials 1–5, the interference conditions (list B and short-delay free recall), cued recall (short and long delay), and delayed recognition (hits and false positives).

Prediction efficacy was first examined using a multiple linear regression analysis in which follow-up CVLT long-delay free recall was regressed against the predictor variables of age, sex, education, baseline long-delay free recall, number of APOE-ε4 alleles, family history of AD (0 = absent, 1 = present), and left-hippocampal volume [19,28,34,49,76,102]. To determine the relative contribution to the variance of each predictor variable, we additionally submitted the data to a stepwise multiple regression analysis entering the same predictor variables using a selection criterion of p < 0.05. We concluded by conducting receiver operating characteristic (ROC) analyses of CVLT decline from baseline to follow-up as predicted by number of APOE-ε4 alleles, left-hippocampal volume, and the combination of APOE-ε4 and left-hippocampal volume, calculated as volume in cm³ minus the number of APOE-ε4 alleles.

3. Results
3.1. Baseline
Raw-score means, least-square means adjusted for the modeled covariates, and standard deviation (S.D.) values of total cerebral-volume measurements are presented in the upper portion of Table 2. Examination of cerebral volume as a function of APOE group revealed a significant main effect of sex (F[1,149] = 63.00, p < 0.0001) but no effect of APOE (F[3,149] = 1.44, p > 0.23) or the APOE × sex interaction (F[3,149] = 1.35, p > 0.25). The same findings were observed combining the APOE-ε3/4 and -ε4/4 groups to increase power: a main effect of sex (F[1,151] = 76.81, p < 0.0001) but no effect of APOE (F[2,151] = 1.73, p > 0.18) or the APOE × sex interaction (F[2,151] = 1.80, p > 0.16).

Results for left- and right-hippocampal volume as a function of APOE group are presented in the lower portion of Table 2. Analyses revealed a significant effect of the cerebral-volume covariate (F[1,140] = 6.95, p < 0.01) and age (F[1,140] = 6.60, p < 0.02) but did not yield a main effect of APOE group (F[3,140] = 0.18, p > 0.91) or an APOE × age interaction (F[3,140] = 0.17, p > 0.91). However, an APOE × age × laterality interaction was observed (F[3,140] = 2.86, p < 0.04), as well as a cerebral volume × laterality interaction (F[1,140] = 4.81, p < 0.03) and a marginally significant APOE × laterality interaction (F[3,140] = 2.64, p < 0.06). No other main effects or interactions approached significance. Combining the APOE-ε3/4 and -ε4/4 groups to increase power yielded identical findings except for a reduction of the APOE × age × laterality interaction to borderline significance (F[2,142] = 2.89, p < 0.06).
Significant correlations were observed between cerebral volume and CVLT trial 5 (−0.23, \(p<0.005\)), total words for trials 1–5 (−0.27, \(p<0.001\)), list B (−0.21, \(p<0.01\)), short-delay free recall (−0.21, \(p<0.01\)), short-delay cued recall (−0.23, \(p<0.005\)), long-delay free recall (−0.21, \(p<0.01\)), long-delay cued recall (−0.23, \(p<0.005\)), and recognition false positives (−0.17, \(p<0.05\)). As the presence of negative correlations suggested that sex effects might have influenced the findings, statistics were repeated with a GLM analysis of the primary measure, long-delay free recall, modeled by number of APOE-\(e4\) alleles, sex, cerebral volume, and the APOE \(\times\) sex interaction. These analyses revealed that cerebral volume remained significant with a negative coefficient (\(F_{[1,150]}=4.68, p<0.04\)), and furthermore a main effect of APOE (\(F_{[1,150]}=4.40, p<0.04\)) and an APOE \(\times\) cerebral volume interaction (\(F_{[1,150]}=4.19, p<0.05\)) were observed. The main effect of APOE derived from poorer performance for \(e4\) homozygotes (7.9 ± 3.6 words) compared with \(e4\) heterozygotes (10.4 ± 3.0 words) and non-\(e4\) subjects (9.9 ± 3.1 words); the interaction with APOE followed from a significant relationship between recall performance and cerebral volume for individuals with no \(e4\) alleles (\(r=-0.30, p<0.005\)) but not for those with at least one \(e4\) allele (\(p>0.66\)). No other analyses involving cerebral volume were significant, and no relationships were observed between the left or right hippocampus and any of the CVLT variables examined.

### 3.2. Follow-up

Raw-score means, covariate-adjusted least-square means, S.D. values, and raw-score ranges for baseline and follow-up CVLT administrations are presented in Table 3. As hypothesized, an interaction for long-delay free recall was observed between APOE \(\times\) time (\(F_{[3,79]}=4.17, p<0.009\)). Tukey pairwise comparisons revealed significant differences between the means of the APOE-\(e3/4\) and -\(e3/3\) groups at follow-up (\(p<0.05\)). Near-significant effects were also observed for sex (\(F_{[1,79]}=3.39, p<0.07\)) and time (\(F_{[1,79]}=3.00, p<0.09\)). Examination of secondary measures revealed an APOE \(\times\) time interaction in the analysis of learning performance from trial 1 to trial 5 (\(F_{[3,79]}=3.02, p<0.04\)). The effect of trials was significant in this analysis (\(F_{[1,79]}=33.12, p<0.0001\)), as well as the effect of sex (\(F_{[1,79]}=7.80, p<0.007\)). Tyuke pairwise comparisons revealed a significant difference between APOE-\(e3/4\) and -\(e3/3\) subjects on trial 5 at follow-up (\(p<0.05\)). A significant APOE \(\times\) time interaction was also observed for total words for trials 1–5 (\(F_{[3,79]}=3.56, p<0.02\)), accompanied by a significant effect of sex (\(F_{[1,78]}=10.68, p<0.002\)) and significant differences between the APOE-\(e3/4\) and -\(e3/3\) groups at follow-up (\(p<0.05\)). An APOE \(\times\) time interaction was noted for long-delay cued recall (\(F_{[3,78]}=3.96, p<0.02\)), with a near-significant effect of sex (\(F_{[1,78]}=3.08, p<0.09\)), but no pairwise comparisons were significant. No significant effects were observed for the other secondary memory measures. Combining APOE-\(e3/4\) and -\(e4/4\) groups to increase power yielded largely identical findings, in addition to an APOE \(\times\) time interaction for short-delay free recall (\(F_{[2,80]}=3.33, p<0.05\), reflecting worse performance at follow-up for APOE-\(e4\) carriers compared with the APOE-\(e3/3\) group (\(p<0.05\)).

### 3.3. Prediction efficacy

Multiple linear regression analysis predicting endpoint CVLT long-delay free recall revealed a total \(R^2\) of 0.56, which was highly significant (\(F_{[7,74]}=13.46, p<0.0001\)). Baseline CVLT recall performance (\(t_{[1]}=7.61, p<0.0001\), APOE-\(e4\) load (\(t_{[1]}=-2.80, p<0.007\)), and left-hippocampal volume (\(t_{[1]}=2.35, p<0.03\)) extracted significant variance. Age, sex, education, and positive family history of AD did not achieve significance in the model.

Forward stepwise multiple regression analysis first selected baseline long-delay free recall (\(F_{[1,80]}=59.26, p<0.0001\)) with a resulting \(R^2\) = 0.43, followed by APOE-\(e4\) load (\(F_{[1,79]}=11.05, p<0.002; R^2=0.50\)), followed by left-hippocampal volume (\(F_{[1,78]}=5.23, p<0.03; R^2=0.53\)). No other variables contributed to the solution.

Fifty-one subjects (61%) declined by at least one word on CVLT long-delay free-recall scores from baseline to follow-up; 33 subjects (39%) achieved identical scores or improved. ROC curves examining decline predicted by APOE-\(e4\) load, hippocampal volume, and the combination of the two are presented in Figs. 1–3, respectively. For APOE-\(e4\) load, the area under the curve (AUC) was 0.61; for hippocampal volume
| Table 3 |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | APOE e2/3        | e3/3            | e3/4            | e4/4            |
| **Trial 1**    |                 |                 |                 |                 |
| **Trial 5**    |                 |                 |                 |                 |
| **Total words (trials 1–5)** |                 |                 |                 |                 |
| Baseline       | 50.6 [47.9] ± 12.1 (30–72) | 49.3 [48.5] ± 11.8 (26–71) | 48.3 [49.5] ± 8.9 (28–70) | 46.7 [49.6] ± 10.0 (17–65) |
| Follow-up      | 47.8 [45.9] ± 6.8 (38–57) | 48.8 [48.2] ± 10.6 (13–67) | 41.3 [42.1] ± 10.7* (18–60) | 38.0 [40.0] ± 10.5 (19–47) |
| **List B**     |                 |                 |                 |                 |
| Baseline       | 6.1 [5.8] ± 1.1 (2–8)  | 6.8 [6.7] ± 2.4 (3–13)  | 6.0 [6.2] ± 1.8 (2–10)  | 5.3 [5.7] ± 0.8 (4–9)  |
| Follow-up      | 6.1 [6.1] ± 1.8 (4–8)  | 5.6 [5.6] ± 2.0 (1–9)  | 4.7 [4.7] ± 1.4 (2–7)  | 4.2 [4.2] ± 2.9 (0–8)  |
| **Short delay**|                 |                 |                 |                 |
| Cued Baseline  | 9.0 [8.5] ± 2.3 (6–13)  | 9.7 [9.5] ± 3.1 (0–15)  | 7.4 [7.6] ± 3.6 (0–15)  | 6.7 [7.2] ± 4.1 (0–11)  |
| **Cued**       |                 |                 |                 |                 |
| Follow-up      | 10.6 [10.5] ± 3.0 (6–15)  | 10.3 [10.3] ± 3.2 (1–16)  | 8.9 [9.0] ± 2.8 (2–14)  | 8.5 [8.6] ± 3.3 (3–13)  |
| **Long delay** |                 |                 |                 |                 |
| Cued Baseline  | 9.6 [9.1] ± 2.3 (7–13)  | 10.1 [10.0] ± 3.5 (0–16) | 7.7 [8.0] ± 3.9* (0–15) | 6.8 [7.4] ± 4.4 (0–12) |
| **Cued**       |                 |                 |                 |                 |
| Baseline       | 11.6 [11.2] ± 2.7 (6–16)  | 10.9 [10.7] ± 3.2 (4–16) | 10.7 [10.9] ± 2.7 (5–16) | 10.0 [10.5] ± 3.5 (0–15) |
| Follow-up      | 11.4 [11.0] ± 2.5 (8–15)  | 10.9 [10.8] ± 3.3 (1–16)  | 8.9 [9.1] ± 3.3 (2–15)  | 7.2 [7.6] ± 4.5 (0–12)  |
| **Recognition**|                 |                 |                 |                 |
| False positives|                 |                 |                 |                 |
| Baseline       | 1.4 [1.5] ± 2.1 (0–5)  | 1.3 [1.3] ± 1.8 (0–9)  | 2.5 [2.5] ± 4.0 (0–21)  | 3.2 [3.0] ± 2.9 (0–13)  |
| Follow-up      | 2.1 [2.5] ± 1.7 (0–4)  | 2.3 [2.4] ± 2.8 (0–14)  | 3.6 [3.5] ± 4.1 (0–18)  | 6.2 [5.8] ± 3.1 (2–9)  |

Data are presented as unadjusted mean [adjusted mean] ± S.D. followed by the range given in parentheses. Adjusted means are least-square means controlling for the sex covariate.

* p<0.05 compared with the APOE-e3/3 group.

Fig. 1. ROC curve of memory decline predicted by number of APOE-e4 alleles.

Fig. 2. ROC curve of memory decline predicted by left-hippocampal volume.
it was 0.55. For the combined APOE-hippocampal measure, values ranged from 1.18 to 5.07, with AUC equaling 0.64.

4. Discussion

This study is one of the first to examine the relative contribution of APOE-ε4 and hippocampal volume to prediction of longitudinal memory performance defined quantitatively in healthy elderly adults. Our findings additionally contribute to the literature on cross-sectional hippocampal-volume differences as a function of APOE-ε4 genotype, as well as the relationship between brain morphometric indices and memory performance as moderated by APOE-ε4. We first discuss cross-sectional followed by longitudinal findings.

4.1. Cross-sectional findings

We found no cross-sectional differences in cerebral or hippocampal volume as a function of APOE genotype in our cohort of healthy controls. Our findings are thus contrary to the report of Yasuda et al. [103], who observed an increase in brain volume with an increasing number of APOE-ε4 alleles. Findings of equal volume across groups for hippocampus, however, are consistent with a number of prior studies [5,63,77,84]. In addition, we observed interactions involving APOE: a significant APOE × age × laterality effect and a near-significant APOE × laterality effect. Examination of the interaction with age and laterality suggested stronger negative age × hippocampus relationships for the left hemisphere in APOE homozygotes (ε3/3 and ε4/4) and stronger relationships for the right in heterozygotes. The near-significant interaction with laterality suggested that the size of the right hippocampus was nearly equivalent to the left hippocampus in the APOE-ε3/4 group but substantially larger in the APOE-ε2/3 and -ε4/4 groups and intermediate in the -ε3/3 group. This suggestive trend is interesting in light of the observation that AAMI individuals had reduced right-greater-than-left hippocampal asymmetry [91] and the continuing controversy over the possible lateralizing effects of APOE isoforms on hippocampal volume [1,3,21,30,35,54,55,73,87,90,96]. Geroldi et al. [29], for example, reported that this asymmetry was reduced with increasing dose of APOE-ε4. We found decreased asymmetry for the APOE-ε3/4 group compared with APOE-ε3/3 subjects and especially the APOE-ε2/3 group. However, the gene-dose effect for ε4 was not observed, as APOE-ε4/4 subjects maintained a high level of the normal right-greater-than-left asymmetry [62,71]. The relationship between APOE genotype and morphometric indices appears to be a complex one, with regionally specific alterations for a given allele combination [26,87] and age-dependent effects [7,13,92]. Further work with larger samples and a greater age range is required to elucidate laterality effects as a function of APOE isoform and the conditions under which homo- versus heterozygosity becomes relevant. More work will also be required to elucidate the underlying mechanisms by which APOE-ε4 might exert differential morphometric changes, depending on allele dose.

We additionally observed a negative cross-sectional relationship between delayed-memory performance and cerebral volume, findings opposite in direction from Kramer et al. [49], who reported a positive relationship between cortical gray-matter volume and immediate free recall, and Tisserand et al. [95], who found that larger brain volume was correlated with superior learning and delayed recall. Although the direction of our findings suggested that they may have been influenced by the well established sex effect on CVLT performance [14], GLM analysis controlling for sex still yielded a significant negative relationship between cerebral volume and delayed recall, as well as a main effect of APOE group and an APOE × cerebral volume interaction. Findings were thus consistent with the observation that larger structural volume is not invariably linked to optimal cognitive and behavioral outcomes [28,48,83,98,99]. We also did not observe a relationship between CVLT performance and volume of either the left or right hippocampus. A number of studies in AD, mild cognitive impairment (MCI), and normal controls have reported an association between delayed-memory performance and hippocampal volume [17,19,31,33,34,48,49,58,68,91]. However, other studies have failed to confirm such results [95,99,104]. We furthermore observed a significant negative relationship between hippocampal volume and age, consistent with some [65,74,86,95] but not all [51,93,99] prior studies.

4.2. Longitudinal findings

Upon following subjects longitudinally for 5 years, we observed a significant decline in CVLT performance in individuals with APOE-ε4, as hypothesized. These findings are consistent with longitudinal declines in performance that have been reported in association with APOE-ε4 [2,6,20,23,36,38,44,61,88,89] and support the sensitivity of delayed-memory impairment to the development of AD [8,16,28,47,48,56,100]. As with our hippocampal findings,
however, results were not straightforward, as pairwise comparisons showed differences between subjects with APOE-\textit{e3/4} versus \textit{e3/3} but not between \textit{e4/4} and \textit{e3/3} or any of the other groups. The mean for the APOE-\textit{e4/4} group was lower than for the APOE-\textit{e3/4} subjects; however, the S.D. for APOE-\textit{e4/4} subjects was higher, thus contributing to this result. Caselli et al. [13] similarly reported lower complex-figure recall for APOE-\textit{e3/4} subjects compared with non-carriers, but no differences were observed between APOE-\textit{e4/4} subjects and non-carriers. The authors also found a significant negative correlation between digit span and age in APOE-\textit{e4/4} subjects that was not observed in non-\textit{e4} individuals. Dal Forno et al. [18] reported that AD patients with APOE-\textit{e3/4} showed more precipitous declines in general mental status and verbal fluency than non-carriers and \textit{e4} homozygotes. Findings thus suggest that individuals with just one APOE-\textit{e4} allele show brain changes that are unique from other APOE genotypes.

We additionally observed an APOE-\textit{e4}-associated longitudinal decline in learning efficiency as defined by the normal increase in performance from CVLT trial 1 to trial 5. These data are consistent with those reported by Lange et al. [52]. Such findings reflecting acquisition and encoding deficiencies suggest that brain regions in addition to mesial-temporal areas may be affected in APOE-\textit{e4} individuals and is consistent with investigators who argue that cognitive functions other than mnemonic operations are compromised with APOE-\textit{e4}, particularly prefrontal-executive attentional mechanisms [13,56,81,82,101]. Findings do not, however, agree with Nilsson et al. [67], who reported that APOE-\textit{e4} exerted its strongest effect in the presence of memory-retrieval support. We found free-recall and learning measures to be more sensitive to APOE-\textit{e4} effects than cued recall, and no differences were observed for recognition testing, which provided the greatest amount of retrieval support.

4.3. Prediction efficacy

We observed in multiple linear regression analysis that CVLT recall performance after a delay interval was accounted for most strongly by score on the task 5 years earlier, followed by number of APOE-\textit{e4} alleles, followed by left-hippocampal volume. The relative magnitude of the findings suggests that prior ability or background cognitive reserve is the predominant influence in predicting performance over time, followed by the genetic contribution of APOE-\textit{e4} load and structural brain characteristics independent of APOE effects. These variables were stronger predictors than age, sex, educational level, and family history of AD. The findings are therefore consistent with our second hypothesis and with studies showing that baseline hippocampal volume predicted memory decline over time or that hippocampal volume discriminated between normals and those with cognitive impairment [43,51,102]. Our findings are also consistent with Lange et al. [52], who examined conversion to AD from normal and preclinical-AD classifications in a 2-year longitudinal study and found that APOE-\textit{e4} was a significant predictor in forward stepwise multiple regression models examining age, sex, education, and eight variables from the CVLT and Wechsler Memory Scale. Findings are in contrast, however, to Dal Forno et al. [18], who used a backward stepwise regression analysis in AD patients and found that APOE was not significant in predicting rate of decline when considered together with handedness, education, family history of dementia, and performance on five neuropsychological tests. Convit et al. [16] found that MCI subjects who converted to AD over 3.2 years had smaller hippocampus at baseline. They also found in a stepwise backward logistic regression model that immediate paragraph recall and MMSE score were retained, whereas delayed memory and digit span were not. None of the cognitive measures, however, improved classification accuracy after accounting for morphometric indices, and APOE-\textit{e4} was not associated with decline. Jack et al. [41] also examined MCI patients at baseline and examined conversion to AD after an average of 2.7 years. Only hippocampal volume, age, and two of seven cognitive measures (MMSE and verbal fluency) were significant predictor variables for crossover, and only a trend was found showing an increased relative risk for APOE-\textit{e4}-positive individuals. Our study is unique from each of these prior investigations in using a sample composed entirely of normal individuals at baseline and its focus on objective, repeated testing of memory performance rather than clinical diagnosis of conversion to AD.

Consistent with findings for regression analyses, ROC curves predicting decline in CVLT performance over the 5-year follow-up period revealed greater AUC for number of APOE-\textit{e4} alleles compared with left-hippocampal volume. Optimal AUC was obtained by combining indices, i.e., by subtracting the number of APOE-\textit{e4} alleles from hippocampal volume in cm\textsuperscript{3}. As such, findings from regression and ROC analyses supported our third hypothesis that the combination of APOE-\textit{e4} and hippocampal volume would better predict follow-up memory performance than APOE-\textit{e4} alone. Only 56% of the explained variation was accounted for by all variables, however, and the increment in variance contributed by hippocampus over APOE in stepwise analysis was only 3%. Similarly, all AUC statistics were poor in their overall magnitude, and the most promising measure—the combination of APOE-\textit{e4} load and left-hippocampal volume—was a novel one. Further work would be required to establish this combined index as viable. Additional variables predicting longitudinal memory decline in elderly cohorts must also be identified.

4.4. Limitations and conclusions

Limitations of our study include the selection of many subjects from an aging-center subject registry, who tended to have higher educational achievement than the population at large. This recruiting decision may have biased the sam-
ple in favor of increased cognitive reserve. Participants may have also been especially motivated individuals in agreeing to come back to our medical center for a longitudinal evaluation, or they may have been concerned about potential memory loss. Our results thus might not generalize to a random community sample with a lower level of education or interest in volunteer time commitments. Sex compositions were exaggerated in favor of women in the APOE-ε2/3 and -ε3/3 groups and men in the APOE-ε4 groups, although statistical analyses controlled for sex. We had only nine APOE-ε4/4 subjects at baseline and six at follow-up, and thus inferences regarding this group must be tentative. Analyses performed after combining the APOE-ε3/4 and -ε4/4 groups, however, did not change the fundamental findings. We only evaluated verbal memory and thus cannot infer about visual memory or other cognitive functions [13]. Handedness data were also not recorded at baseline. We had only one analyst measuring hippocampus, and thus while intra-rater reliability was high, we could not establish inter-rater reliability. As with most studies of this kind, some of the subjects may have been manifesting incipient dementia, and thus some participants may have been misclassified. In the follow-up sample, four subjects were entirely amnesic on long-delay free recall and exhibited testing protocols that were ominous for dementia: one had APOE-ε3/3; two had APOE-ε3/4, and one had APOE-ε4/4. Subjects did not undergo a formal diagnostic work-up, however, nor are autopsy data available. The small number of amnesic subjects and the approximately equal representation proportionally across APOE groups would indicate that they did not exert a disproportionate influence on the findings.

Another limitation of the research includes the fact that it was originally designed with a 3-year follow-up interval but was subsequently modified into a 5-year study. For this reason, we experienced a greater-than-usual attrition rate. However, individuals with repeat examinations were not significantly different from those lost to follow-up on APOE classification, age, sex, education, MMSE score, family history of AD, and 9 of 10 CVLT variables. This predominant lack of difference between samples adds confidence to our findings. The only measure to differ was list B, a measure of proactive memory interference. Interestingly, this was the only CVLT measure observed in our prior study to differ between subjects with a positive versus negative family history of AD [14]. It is possible that this measure may be associated with increased morbidity for reasons that are presently unclear.

Future studies with larger sample sizes are required to confirm these findings and to explore the relative contribution of APOE-ε4 and hippocampal volume to predicting response to treatments for AD and MCI.

Acknowledgment

This work was supported by 1-K-08-MHO1460.

References


