Hippocampal Atrophy Confounds Template-Based Functional MR Imaging Measures of Hippocampal Activation in Patients with Mild Cognitive Impairment

BACKGROUND AND PURPOSE: Functional MR imaging has been used to study patterns of hippocampal activation that distinguish pathologic from normal memory loss in the elderly population. Our objective was to assess whether hippocampal atrophy confounds measurements of hippocampal activation in subjects with mild cognitive impairment (MCI).

METHODS: Twenty subjects with MCI and 20 elderly control subjects with objectively normal memory were studied at 4T during a face-name paradigm designed to activate the hippocampus. Hippocampal activation was measured using 2 separate approaches: applying a preset region of interest (ROI) in standardized template space and applying a manually drawn ROI in native subject space. Pearson correlation coefficients were calculated to compare group-dependent relationships between hippocampal volume and activation. Analysis of covariance (ANCOVA) was performed to assess group differences in hippocampal activation during encoding and retrieval. Age and hippocampal volume were included as covariates, as was a term for the interaction between hippocampal volume and group.

RESULTS: When hippocampal activation was measured by the template-based method, the correlation coefficient in the right hippocampus of subjects with MCI but not control subjects during retrieval differed significantly from zero. There was a significant (P < .05) group-by-volume interaction in the ANCOVA model. No significant correlations or interactions were demonstrated when activation was measured in native subject space with manually drawn ROIs.

CONCLUSION: Our findings suggest a potential confounding relationship between hippocampal volume and activation for subjects with MCI in template-based analyses. Template-based measures of hippocampal activation that do not adequately account for hippocampal atrophy should be used with caution in patients with MCI.
standardized template methods based upon a whole-brain stereotactic space. A previous fMRI study that used a template-based analysis technique demonstrated a correlation between atrophy and activation in the inferior frontal gyrus of patients with AD but not control subjects during a semantic decision task. Recent studies have instead used a manually drawn region-of-interest (ROI) approach, which intrinsically takes into account individual differences in hippocampal volume and probably yields the most accurate measurement of hippocampal activation independent of volume. This technique, however, is time-consuming, labor-intensive, and has uncertain reproducibility across individual operators and institutions.

The purpose of this study was to assess whether hippocampal atrophy in the MCI population confounds fMRI measures of hippocampal activation that use template-based ROIs. To answer this question, we studied the relationship between hippocampal volume and activation using both template-based and manually drawn ROI methods in a group of subjects with MCI and elderly control subjects. We hypothesize that template-based ROI analysis methods of assessing fMRI activation will demonstrate a correlation with brain atrophy in the MCI group, suggesting that atrophy is a confounding factor in measurements of hippocampal activation. Further, we hypothesize that ROI-based analysis, which accounts for changes in hippocampal volume, will not be significantly affected by hippocampal atrophy.

**Subjects and Methods**

**Subjects**
The study received approval from our Institutional Review Board. Subjects satisfying entry criteria were recruited from the local community via advertisements and referrals. Twenty subjects with MCI and 20 cognitively normal elderly subjects were studied.

**Entry Criteria**
Requirements for entry included fluency in English, completion of at least 8 years of formal education, and willingness to participate in a functional MR imaging scan. All subjects provided written informed consent before any testing or neuropsychologic evaluation. Subjects satisfying entry criteria were recruited from the local community via advertisements and referrals. Twenty subjects with MCI and 20 cognitively normal elderly subjects were studied.

MCI. Inclusion criteria were as follows for the MCI group: (1) recent history of symptomatic worsening in memory supported by informant, (2) objective memory impairment (>1 SD below normal) as evidenced by performance on the California Verbal Learning Test (CVLT) II and logical memory and visual reproduction tests from the Wechsler Memory Scale III, (3) normal/near-normal performance on global cognitive test, defined as a Mini-Mental State Examination (MMSE) score >24, (4) global rating on the Clinical Dementia Rating (CDR) scale of 0.5 (questionable dementia) with at least 0.5 on the memory score, (5) does not meet National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) or Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) criteria for dementia, (6) normal/near-normal independent function, and (7) absence of other factors that might better explain memory loss (eg, current major depression).

Control Subjects. Criteria were as follows for inclusion in the control group: (1) MMSE score >28, (2) does not meet NINCDS-ADRDA or DSM-IV criteria for dementia, (3) normal/near-normal independent function, (4) normal memory, and (5) CDR global score of 0.

**Exclusion Criteria**
Subjects were excluded on the basis of the following: (1) uncontrolled depression or other psychiatric illness, (2) taking psychoactive medications known to significantly affect memory, (3) standard contraindications to MR imaging, including size incompatible with scanner, metal implants, or cardiac pacemakers, (4) technical difficulties that prevented the completion of successful anatomic imaging and all functional MR imaging runs, and (5) excessive motion during the functional MR imaging examination in excess of 5 mm determined by center of mass plots.

**fMRI Stimuli and Paradigm**
Our study used a variation of a face-name associative memory encoding task developed by Sperling et al in young adults and later applied to elderly subjects and patients with AD. In addition to the encoding task, our paradigm also included a retrieval task, for which responses were monitored within the scanner. Subjects were required to encode and later retrieve face-name associations of 2 conditions, novel face-name pairs and familiar face-name pairs, which were presented to the subject within a blocked experimental design. Sixty novel and 2 familiar face-name pairs drawn from the AR Face Database were presented within a block design over a period of 6 minutes 50 seconds per run for a total of 3 runs.

**Anatomic and Functional Whole-Brain Imaging**
Imaging was performed on a 4T MR scanner. Axial T2-weighted spin-echo images (matrix, 256 × 256; 3.75-mm sections with no intersection gap; field of view (FOV), 240 mm) were obtained through the brain for diagnostic purposes to assess for intracranial pathology. Anatomic scans, consisting of 44 contiguous 3.75-mm coronal sections, were acquired for normalizations of the functional scans using high-resolution T1-weighted images (IR prep3D SPGR, echo time [TE], 5.4 ms; repetition time [TR], 12.2 ms; inversion time [TI], 500 ms; flip angle [FA], 20°; matrix, 256 × 256; FOV, 240 mm). Functional scans, consisting of a time series of 164 T2*-weighted isotropic image volumes (inverse spiral echo-planar imaging; TE, 31 ms; TR, 2500 ms; FA, 60°; matrix, 64 × 64; FOV, 240 mm), were acquired during each of the 3 functional runs per subject from the same 44 continuous anatomic series coronal section locations.

All functional images were initially screened for quality control purposes using center-of-mass plots to detect excessive motion or section acquisition errors. Anatomic images were screened for intracranial pathology by a board-certified neuroradiologist (J.R.P.).

**Hippocampal Volume Measurements**
For hippocampal tracing, T1-weighted images were viewed using ITK-SNAP 1.0 software, a computer program that displays structural images simultaneously in 3 planes and allows for mouse-driven manual segmentation of images. The left and right hippocampi were traced independently by a single experimenter (S.K.) who was blinded to clinical classification. Tracings were similar to those reported in a previous volumetric study. The hippocampus was first identified at the most posterior portion, where the hippocampal tail became visible under the fornix, and proceeded on contiguous sections to the most anterior section on which the hippocampus could be delineated from the amygdala. The selected area was then reviewed for consistency on the axial and sagittal presentations. The traced region en-
which were obtained from anatomic MR images in the same orientation as the subject’s contrast map to produce the activated voxel count within the hippocampus in native subject space.

After manual tracing was completed, the volume within the encompassed hippocampal region was calculated for right (RHV) and left (LHV) sides. Hippocampal volumes were then adjusted to account for individual differences in intracranial volume. As an index of total intracranial volume, single-section intracranial volume was reported by Laakso et al. for the adjustment of hippocampal volume. Applying this method, intracranial volume (ICA) was measured on the coronal section at the level of the anterior commissure. Right and left hippocampal volumes were divided by this measure of intracranial volume resulting in RHV/ICA and LHV/ICA, respectively.

**Hippocampal ROI Analysis in Template-Based Space**
Processing of images was performed using Statistical Parametric Mapping 2 (SPM2) software. For the template-based method, pre-processing of each subject’s data consisted of section timing and motion correction, normalization to the MNI template, and spatial smoothing with an 8-mm Gaussian kernel. Two contrast maps of the novel versus familiar condition were created for each subject using the general linear model approach in SPM2, one for encoding and one for retrieval. Right and left hippocampi were analyzed independently for each part of the memory task (encoding and retrieval). Statistical t-maps were created for each subject using a first-level analysis with a voxel-wise significance level of $P < .01$ (uncorrected) for all maps. This threshold was selected for group comparisons consistent with previously published fMRI studies in similar populations. The number of activated voxels was measured within the standardized hippocampal ROI (Fig 1A) defined by WFU PickAtlas. The proportion of activated voxels was calculated for each subject by dividing the number of activated voxels by the number of voxels in the standardized hippocampal ROI (922 and 946 voxels for left and right hippocampi, respectively).

**Hippocampal ROI Analysis in Native Subject Space**
For the analysis within native subject space, processing of each subject’s data was identical to that of the template-based analysis with the exception of normalization to the MNI template. Before spatial smoothing, the data were resampled to match the voxel size of the standard MNI space ($2 \times 2 \times 2$ mm). The manually drawn ROIs, which were obtained from anatomic MR images in the same orientation as those of the functional data (Fig 1B), were then applied to each subject’s contrast map to produce the activated voxel count within each ROI. The voxel count was then adjusted for hippocampal volume by dividing activated voxels by voxel count within each ROI. This yielded the proportion of activated voxels within the subject’s manually drawn ROIs.

**Table 1: Demographic and clinical characteristics**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Mild Cognitive Impairment ($N = 20$)</th>
<th>Control ($N = 20$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, $\bar{y}$ (SD)</td>
<td>75.0 (7.6)</td>
<td>71.2 (4.5)</td>
</tr>
<tr>
<td>Age range, $\bar{y}$</td>
<td>55.5–83.3</td>
<td>63.3–80.5</td>
</tr>
<tr>
<td>Men/women</td>
<td>12/8</td>
<td>9/11</td>
</tr>
<tr>
<td>Education, $\bar{y}$</td>
<td>15.0 (2.2)</td>
<td>15.9 (2.9)</td>
</tr>
<tr>
<td>CVLT (SD)</td>
<td>5.0 (2.5)</td>
<td>11.0 (2.6)*</td>
</tr>
<tr>
<td>MMSE (SD)</td>
<td>28.7 (1.5)</td>
<td>28.4 (1.4)*</td>
</tr>
<tr>
<td>LHV (as % of ICA) (SD)</td>
<td>6.19 (1.16)</td>
<td>7.29 (1.01)*</td>
</tr>
<tr>
<td>RHV (as % of ICA) (SD)</td>
<td>6.40 (1.00)</td>
<td>7.16 (0.91)*</td>
</tr>
</tbody>
</table>

**Statistical Analysis**
The 2 groups were compared for differences in demographic and clinical variables using 2-tailed Student $t$ tests (for age, education, MMSE, CVLT, and hippocampal volume) or $\chi^2$ analysis (for sex). Pearson correlation coefficients were calculated to directly compare the group-dependent relationships between hippocampal volume and activation. Likewise, Pearson correlation coefficients were used to examine correlations between hippocampal activation and clinical data of MMSE and CVLT. Analysis of covariance (ANCOVA) was performed with hippocampal activation during encoding or retrieval as the independent variable, group as the main effect, and age and hippocampal volume as covariates. The model also included an interaction term between hippocampal volume and group.

**Results**
Sex and education did not differ significantly between the 20 subjects with MCI and the 20 control subjects (Table 1). The MCI group was marginally older than the control group ($P = .056$), so age was accounted for in the statistical model. As expected, CVLT and MMSE scores were significantly lower in subjects with MCI ($P < .001$). There was no significant correlation between hippocampal activation detected with a manual ROI and clinical data, including CVLT or MMSE. After adjusting for differences in intracranial volume, RHV and LHV were significantly smaller in subjects with MCI, with a greater difference in mean volume between the groups noted on the left.

In subjects with MCI, a significant negative correlation between hippocampal volume and activation ($r = -0.477$, $P < .05$) was noted in the right hippocampus during retrieval; this correlation was detected only in template-based analysis. Native space analysis with manually drawn ROIs failed to reveal a significant relationship ($r = 0.204$) in subjects with MCI. Conversely, correlations in control subjects did not differ significantly from zero or from each other, regardless of the method of analysis used for hippocampal activation. These group-specific relationships in the right hippocampus during retrieval are shown in Fig 2 (template-based analysis with automated hippocampal ROI) and Fig 3 (native subject space with man-
The relationship between volume and activation was represented as proportion of ROI activated. The line demonstrates a significant negative correlation between hippocampal volume and activation in the subjects with MCI.

In the ANCOVA model, there was a significant effect of both group (P = .030) and the interaction term (P = .020) only on right hippocampal activation during retrieval using template-based analysis (Table 2). Native space analysis failed to show a significant effect of group or of the interaction term. Thus, when the effects of atrophy were accounted for by manually drawing the hippocampal ROI and performing analysis in native space, no significant difference was identified in the amount of activation between control subjects and those with MCI, as evidenced by lack of a statistically significant group effect or interaction term.

### Discussion

Template-based analyses offer speed, reproducibility, and operator independence and thus have wide appeal, particularly for use in whole-brain comparisons of functional activation. Whole-brain, template-based analyses allow detection of patterns of cortical function and dysfunction throughout the entire brain in an exploratory analysis or within one or multiple standardized ROIs in a hypothesis-driven analysis. The operating assumption is that spatial normalization corrects for differences in size of various cortical structures. However, this might not be true for localized areas of atrophy; in fact, the use of standardized ROI analysis in the MTL of memory-impaired subjects has been questioned in the past. When localized atrophy is a concern, such as in subjects with MCI, manually drawn ROIs in native subject are frequently used, because they directly account for individual differences in hippocampal volume. We report a significant negative correlation between hippocampal volume and hippocampal activation in a sample of subjects with MCI using a standardized ROI approach in template space, but not a manually drawn ROI approach in native subject space. This suggests that atrophy in a memory-impaired population may act as a confounder in template-based approaches, despite their obvious logistic advantages.

The implication is that reports of areas of significant activation and nonactivation might be spurious when failing to account for differences in volume in structurally and functionally complex regions, such as the hippocampus. A negative correlation between hippocampal volume and hippocampal activation was only evident in subjects with MCI during template-based analysis. This would suggest that subjects with MCI with high hippocampal volumes tend to have much lower levels of activation during retrieval, whereas subjects with MCI with small RHVs have increased levels of hippocampal activation, indicating compensatory right hippocampal activation. Although this hypothesis has been suggested previously in the pathologic progression of MCI, we propose that in template-based studies, this is, at least in part, a confounding rather than real effect. For example, with the

### Table 2: Results from ANCOVA (with P values for independent variables included in the model)

<table>
<thead>
<tr>
<th>ANCOVA Model</th>
<th>Template-Based Analysis</th>
<th>Native Subject Space with Manual ROI</th>
</tr>
</thead>
<tbody>
<tr>
<td>LH, encoding</td>
<td>Group .699 .617</td>
<td>Group .694</td>
</tr>
<tr>
<td>Age</td>
<td>Group .616 **.017</td>
<td>Group .628</td>
</tr>
<tr>
<td>LHV/ICA</td>
<td>Group X LHV/ICA .428</td>
<td>Group X LHV/ICA .700</td>
</tr>
<tr>
<td>RH, encoding</td>
<td>Group .591 .542</td>
<td>Group .593</td>
</tr>
<tr>
<td>Age</td>
<td>Group .271 .181</td>
<td>Group .181</td>
</tr>
<tr>
<td>LHV/ICA</td>
<td>Group X LHV/ICA .494</td>
<td>Group X LHV/ICA .619</td>
</tr>
<tr>
<td>RH, retrieval</td>
<td>Group .030**</td>
<td>Group .030**</td>
</tr>
<tr>
<td>Age</td>
<td>Group .707 .828</td>
<td>Group .828</td>
</tr>
<tr>
<td>RHV/ICA</td>
<td>Group X RHV/ICA .812</td>
<td>Group X RHV/ICA .121</td>
</tr>
<tr>
<td>RH, retrieval</td>
<td>Group .020 .712</td>
<td>Group .020 .712</td>
</tr>
</tbody>
</table>

Note—Dependent variable is proportion of activated voxels in hippocampal region of interest.

LH indicates left hippocampal activation; LHV/ICA, left hippocampal volume divided by single-section intracranial volume; RH, right hippocampal activation; RHV/ICA, right hippocampal volume divided by single-section intracranial volume; Group X LHV/ICA and Group X RHV/ICA, interaction term between group and hippocampal volume. Significant values of P are indicated in boldface type.

*Group term becomes nonsignificant (P = .314) when the interaction term, Group X RHV/ICA, is removed from the ANCOVA model.
current study, a group-specific relationship is no longer evident when using the manually drawn ROI analysis in native subject space, which takes into account individual differences in volume when detecting hippocampal activation. When a standardized template is used to define activation within the hippocampus in subjects with localized hippocampal atrophy, the resulting ROI may spuriously include surrounding tissues, such as the parahippocampal gyrus, and activation within these tissues may account for the apparent increase. In fact, it has been suggested that it may be these extrahippocampal MTL tissues that display compensatory activation to overcome dysfunctional activation within the hippocampus. Unfortunately, our study did not assess activation in these extrahippocampal regions, and we therefore cannot verify this explanation. Nevertheless, it is important to keep in mind that when comparing healthy control subjects with lesser degrees of hippocampal atrophy to memory-impaired subjects with significant hippocampal volume loss, the spurious inclusion of nonhippocampal tissue in template-based ROIs might lead to the false conclusion of increased hippocampal activation in the group with more profound hippocampal atrophy. A similar inverse relationship between local volume loss and activation has been reported in the left inferior frontal gyrus of patients with AD, but not in healthy control subjects, during a semantic decision task. Although the measure of activation and voxel-wise approach differed slightly from our methods, this study also used a template-based approach. Template-based analyses comparing different clinical groups, therefore, could report significant results that might be reflective of morphologic changes rather than effects of neuronal dysfunction in the targeted region.

Manual tracing of hippocampal ROIs provides a direct measure of volume that can be applied to analyses of hippocampal activation to account for atrophy. In this study, we analyzed hippocampal activation within a manually drawn ROI in native subject space that encompassed the full volume of the hippocampus. This approach probably yields the most accurate measurement of hippocampal activation independent of volume but is time-consuming, labor-intensive, and unlikely to be suited to a clinical environment. Other less time-consuming methods of accounting for hippocampal atrophy might be equally valid in the context of template-based analyses. For example, a visual rating scale of hippocampal atrophy (as a gross assessment of CSF accumulation in the perihippocampal fissures) has been used effectively in volumetric studies. Alternatively, Gao et al measured the thinnest width of the MTL at the level of the intercollicular sulcus to differentiate between patients with mild AD and control subjects. Indices of hippocampal volume such as these might be acquired quickly and yet still adequately account for atrophy in template-based fMRI analyses.

It is noteworthy that we were unable to demonstrate a significant correlation between clinical measures of disease and activation, even for the native subject space approach. Other groups have found group differences in activation between subjects with MCI and control subjects with the use of the native space approach. However, the direction of these changes has conflicted, possibly reflecting high-versus low-functioning samples within the MCI population. We were unable to demonstrate a group difference in activation, perhaps reflecting the heterogeneity of our sample of the MCI population and the use of very high field strength at 4T, which may decrease signal intensity-to-noise in areas near the skull base secondary to susceptibility artifacts.

A limitation of this study is that we studied only a single standardized whole-brain template, though the one chosen is commonly used in studies of this nature. We did not specifically investigate the validity of the standardization process itself. It is noteworthy that its validity in a memory-impaired population has been previously addressed. Analysis using probabilistic atlases in disease populations may avoid such confounders, though this complicates comparisons between diseased and nondiseased populations. A second limitation is that different volumes of hippocampal tissue were labeled in the native and template space. A manually drawn ROI was applied to data in native subject space, whereas a predefined anatomic label (WFU PickAtlas Tool) was applied to identify the hippocampus in template space. We cannot comment on which label more accurately identifies hippocampal tissue and excludes surrounding tissue. Nevertheless, we acknowledge that the use of automated spatial normalization and preconfigured anatomic labels has substantial appeal in allowing for standardization across studies. A third limitation is that our findings apply only in the hippocampus. Although there is evidence that template-based measurements of activation may be confounded by atrophy in other areas of the brain in a memory-impaired population, comparisons of template-based and manual ROI approaches were not performed for other cortical structures; thus, we cannot speculate on the effect of atrophy outside the hippocampus.

In conclusion, we have shown, using a template-based ROI analysis approach, that hippocampal atrophy in subjects with MCI confounds measures of hippocampal activation. When comparing functional activation in the hippocampus across different groups of subjects with MCI, such as in a trial of drug efficacy, template-based analyses might be adequate. However, when using a template-based analysis to compare a memory-impaired population and a nondiseased control population, our findings suggest that some index of atrophy should take into account the interaction between local morphologic changes and fMRI measures of activation in the hippocampus.

Acknowledgments

We thank the subjects who participated in this study.

References