Abstract

MRI research examining structural brain atrophy in Alzheimer’s disease (AD) generally focuses on medial temporal and cortical structures, but amyloid and tau deposits also accumulate in the caudate. Here we mapped the 3D profile of caudate atrophy using a surface mapping approach in subjects with AD and mild cognitive impairment (MCI) to identify potential clinical and pathological correlates. 3D surface models of the caudate were automatically extracted from 400 baseline MRI scans (100 AD, 200 MCI, 100 healthy elderly). Compared to controls, caudate volumes were lower in MCI (2.64% left, 4.43% right) and AD (4.74% left, 8.47% right). Caudate atrophy was associated with age, sum-of-boxes and global Clinical Dementia Ratings, Delayed Logical Memory scores, MMSE decline 1 year later, and body mass index. Reduced right (but not left) volume was associated with MCI-to-AD conversion and CSF tau levels. Normal caudate asymmetry (with the right 3.9% larger than left) was lost in AD, suggesting preferential right caudate atrophy. Automated caudate maps may complement other MRI-derived measures of disease burden in AD.

Keywords: Alzheimer’s disease; Mild cognitive impairment; Normal aging; Caudate nucleus; Brain mapping; Surface mapping; MRI; Atrophy; Alzheimer’s Disease Neuroimaging Initiative

Alzheimer’s disease (AD) is the most prevalent form of dementia, affecting over 24 million people worldwide and over 5 million in the USA alone (Jorm et al., 1987). Memory systems of the hippocampus and medial temporal lobes are among the first to be affected by plaque and tangle pathology, with subsequent cortical pathology and widespread neuronal loss leading to impairments in language, attention, orientation, visuospatial, and executive functions. Traditionally, AD research has focused on regions that degenerate earliest, and less attention has been paid to other subcortical brain regions involved in aging and AD (Rubin, 1999). Preferential caudate atrophy is more typical of motor disorders such as Parkinson’s disease (Nakamura et al., 2001), but beta-amyloid and tau pathology also accumulate in the caudate in AD (Rubin, 1999). Preferential caudate atrophy is more typical of motor disorders such as Parkinson’s disease (Nakamura et al., 2001), but beta-amyloid and tau pathology also accumulate in the caudate in AD (Rubin, 1999). Preferential caudate atrophy is more typical of motor disorders such as Parkinson’s disease (Nakamura et al., 2001), but beta-amyloid and tau pathology also accumulate in the caudate in AD (Rubin, 1999). Preferential caudate atrophy is more typical of motor disorders such as Parkinson’s disease (Nakamura et al., 2001), but beta-amyloid and tau pathology also accumulate in the caudate in AD (Rubin, 1999).
target, in conjunction with others, in the search for AD biomarkers in the brain.

Despite some prior volumetric MRI studies, no study has mapped caudate atrophy and evaluated its clinical and pathological correlates in a large population of AD subjects. Here we extracted 3D models of caudate nucleus shape in 400 subjects from the Alzheimer’s Disease Neuroimaging Initiative (ADNI), to test whether caudate atrophy was related to: (1) AD or MCI diagnosis at time of scanning; (2) baseline cognition assessed using the Mini-Mental State Examination (MMSE), clinical dementia rating (CDR), and immediate and delayed recall Wechsler Memory Scale (WMS); (3) future clinical decline (in MMSE, CDR) 1 year after scanning; (4) conversion from MCI to AD in the subsequent 1 or 2 years; (5) CSF biomarkers of AD (amyloid beta, tau, p-tau); and (6) other influential factors such as age and sex, Geriatric Depression (GD) scale, presence of tremor, abnormal gait, AD- and obesity-related ApoE and FTO genotypes, and body mass index (BMI). We also ranked the relative strength of factors associated with caudate atrophy, expecting strongest associations with global cognition, then with CSF markers, and then with risk-modifying genotypes. We also wanted to determine how AD interacts with the naturally occurring asymmetry in caudate anatomy (where the right caudate tends to be about 4% larger than the left), to see if there was asymmetry in the disease process.

1. Methods

1.1. Subjects

We analyzed 400 baseline T1-weighted structural MRI scans from the ADNI public database (www.loni.ucla.edu/ADNI/Data/), along with associated demographic information, ApoE and FTO genotypes, CSF biomarker measures (for amyloid beta, tau, p-tau), and clinical and cognitive assessments (detailed below). Data were downloaded on or before June 1, 2009, and reflect the status of the database at that point. ADNI is a 5-year study launched in 2004 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and nonprofit organizations, as a US$60mn public-private partnership. The primary goal of ADNI has been to test whether serial MRI, PET, biological markers, and clinical and neuropsychological assessments acquired at multiple sites (as in a typical clinical trial), can replicate results from smaller single site studies measuring the progression of MCI and early AD. A major focus of AD biomarker research is the pre-disease state known as mild cognitive impairment or MCI. Individuals with MCI have a 4–6-fold elevated risk for developing AD in the future, with 15% of MCI individuals developing AD within 1 year (Petersen et al., 1993). Determination of sensitive and specific markers for very early AD progression will aid researchers and clinicians in monitoring the effectiveness of new treatments, and lessen the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner MD, VA Medical Center and University of California, San Francisco.

ADNI collects a thorough battery of clinical and cognitive measures. We analyzed several scores that reflect cognitive decline in MCI and AD (Table 1). A total of 100 individuals diagnosed with probable AD (mean age: 75.86 ± 72.5 years), 200 with MCI (75.45 ± 7.03), and 100 healthy elderly controls (76.62 ± 4.83) were studied. The subject pool was age- and sex-matched within each diagnostic group. Data for all 400 subjects were available on each of the clinical measures analyzed in this study, leading to differing sample sizes for some of the following analyses.

Cognitive measures examined included: (1) the Mini-Mental State Examination (MMSE) – a global measure of mental status, evaluating five cognitive domains: orientation, registration, attention and calculation, recall, and language (Cockrell and Folstein, 1988); (2) the global and “sum-of-boxes” clinical dementia rating (CDR and CDR-SB respectively), which measures dementia severity by evaluating patients’ performance in six domains: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care; CDR scores of

<table>
<thead>
<tr>
<th>Table 1 Demographic and clinical data</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Control</td>
</tr>
<tr>
<td>MCI</td>
</tr>
<tr>
<td>AD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Delayed Logical Memory</th>
<th>Immediate Logical Memory</th>
<th>MMSE change at 1 year</th>
<th>Diagnosis change after 1 year (number of AD conversions)</th>
<th>Diagnosis change after 2 years (number of AD conversions)</th>
<th>BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>12.99 (3.23)</td>
<td>11.14 (3.17)</td>
<td>0.14 (1.42)</td>
<td>36</td>
<td>24</td>
</tr>
<tr>
<td>MCI</td>
<td>3.73 (2.79)</td>
<td>6.94 (3.24)</td>
<td>1.21 (2.88)</td>
<td>36</td>
<td>24</td>
</tr>
<tr>
<td>AD</td>
<td>1.11 (1.74)</td>
<td>3.91 (2.82)</td>
<td>2.62 (4.57)</td>
<td>36</td>
<td>24</td>
</tr>
</tbody>
</table>

Mean demographic and clinical data are shown by diagnostic group for all subjects at the time of scanning (baseline); standard deviations are in parentheses. Changes in MMSE scores were measured over a 1-year follow-up interval. Changes in diagnosis were determined after one- and 2-year follow-up intervals; the number of MCI and control subjects who converted to AD is also shown.
the x-, y-, and z-dimensions, yielding a voxel size of $1.25/100^3$ field of view, and a 192 acquisition matrix in $166^2$ acquisition matrix in $1.2$ mm $^3$, later reconstructed to 1 mm $^3$ isotropic.

1.2. MRI acquisition and preprocessing

Scans were acquired on 1.5T MR scanners at 60 sites across the USA and Canada. Different types of scanners (GE, Siemens, or Philips) and various software platforms were used, but a standardized MRI protocol ensured cross-site comparability (Jack et al., 2008). A typical 1.5T MR protocol involved a 3D sagittal MP-RAGE scan with repetition time (TR): 2,400 ms, minimum full excitation time (TE), inversion time (TI): 1,000 ms, flip angle: 8°, 24 cm field of view, and a $192 \times 192 \times 166$ acquisition matrix in the x-, y-, and z-dimensions, yielding a voxel size of $1.25 \times 1.25 \times 1.2$ mm $^3$, later reconstructed to 1 mm $^3$ isotropic voxels.

Image corrections were applied using a processing pipeline at the Mayo Clinic, consisting of: (1) correction of geometric distortion due to gradient nonlinearity (Jovicich et al., 2006), i.e. “grad warp” (2) “B1-correction” to adjust for image intensity inhomogeneity due to B1 nonuniformity (Jack et al., 2008), (3) “N3” bias field correction for reducing residual intensity inhomogeneity (Sled et al., 1998), and (4) geometric scaling to remove scanner- and session-specific calibration errors using a phantom scan acquired for each subject (Gunter et al., 2006). All original image files as well as all corrected images are available at www.loni.ucla.edu/ADNI/Data/.

To adjust for global differences in brain positioning and scale across individuals, all scans were linearly registered to the stereotaxic space defined by the International Consortium for Brain Mapping (ICBM-53) (Mazziotta et al., 2001) with a 9-parameter (9P) transformation (three translations, three rotations, three scales) using the Minctracc algorithm (Collins et al., 1994). This registration is one standard method to account for global differences in brain scale across subjects; alternatives include scaling each subject’s caudate nucleus volume for overall head size, using scalp images, or covarying for total intracranial volume in a general linear model. Globally aligned images were resampled in an isotropic space of 220 voxels along each axis (x, y, and z) with a final voxel size of 1 mm $^3$.

1.3. Automated segmentation

We automatically extracted models of the caudate nucleus from each registered MRI scan, using an automated segmentation method based on adaptive boosting (AdaBoost), that we recently developed and validated (Morra et al., 2008). Adaboost is a machine learning approach that learns to segment a structure in new images based on a small training set of expertly delineated tracings. As in our hippocampal studies (Morra et al., 2009a, 2009b), we created a small representative training set of 21 MRI scans (7 AD, 7 MCI, and 7 controls) that were expertly hand-labeled according to a validated protocol (Looi et al., 2007). These manual traces were then used to produce automated segmentations of the large (non-overlapping) testing set of 400 MRI scans. Using the image intensities and gradients in the training set scans, and statistical information on the likely position and geometry of the caudate, the algorithm learns a classification rule to designate each voxel as caudate or non-caudate. The automated segmentation incorporates ~13,000 features including intensity, combinations of x, y, and z coordinates, and gray matter, white matter, and cerebrospinal fluid tissue classifications. As is standard in machine learning, images used in the training set were excluded from all analyses presented here.

1.4. Volumetric analysis

For each subject, we determined left and right caudate volumes from the automatically generated segmentations. From these, we calculated the mean volume and a measure of asymmetry ($L-R)/(2(L+R))$ for each diagnostic group
and tested for group and hemispheric differences in caudate volume.

1.5. Statistical maps

Parametric surface models were created from the 400 automatically generated caudate segmentations, which were then used to create statistical maps associating local volumetric differences with clinical and cognitive scores (Csernansky et al., 2004; Wang et al., 2007). As shown in Figure 1, surface models were created from each subject’s automatically generated binary segmentation (Thompson et al., 2004). A central curve was calculated through the longitudinal axis of each model; the radial distance from this medial line to each surface point was used as a highly localized measure of atrophy. This same approach was used in two prior papers based on manual tracings (Butters et al., 2008) and tracings from another automated segmentation algorithm (Becker et al., 2006). We have used the same radial distance mapping approach to map hippocampal and ventricular shape differences (Chou et al., 2009; Morra et al., 2009a, 2009b).

Based on a computed point-wise correspondence of the structure surfaces across subjects, geometric surface averaging was performed across all subjects in each diagnostic group. At each surface point, a correlation was run to compare diagnostic groups and determine the association of diagnosis or clinical scores with atrophy, as measured by differences in radial distance. In all maps shown, we used a multiple regression adjusted for age and sex. In these regressions, adjustments were made for age and sex by including them as covariates in the model. We did not adjust our maps for effects of ApoE genotype because we did not detect any significant effect of ApoE genotype on caudate nucleus shape in either the pooled sample or in any diagnostic group, considered separately. All maps presented in this manuscript refer to one-sided tests in the direction

Fig. 1. Mapping caudate atrophy. After image preprocessing, our Adaboost algorithm automatically segments the caudate nucleus bilaterally. A central curve (top right) was calculated through the longitudinal axis of each caudate and the radial distance to each point in a 3D surface mesh was used as a highly localized measure of atrophy (middle right). p-values are calculated at each surface point (bottom right) showing the significance of differences in radial distances between diagnostic groups or their associations with clinical scores. These maps visualized the profile of local shape differences or their clinical correlates at a point-wise level.
expected for the values of each clinical measure. For example, a lower MMSE score indicates poorer cognitive performance, which is expected to correlate with caudate nucleus atrophy rather than expansion. For a measure such as MMSE, we used a positive one-sided test, corresponding to the hypothesis that a lower MMSE score (poorer performance) would be associated with volume reductions in the caudate nucleus. Negative one-sided tests were used for measures in which higher scores are expected with greater atrophy in the caudate, such as BMI, in which a lower score is generally considered indicative of better health.

Color coded *p*-values were mapped onto the average left and right caudate models. To correct for multiple comparisons across surface points, permutation tests provided an overall significance value for each statistical map. Specifically, the suprathreshold area (points with *p* < 0.01) in the map was compared with a null distribution for this same area, estimated from 100,000 random permutations of the covariates (Thompson et al., 2003).

2. Results

2.1. Volumetric results

First, we performed a volumetric analysis of the left and right caudate nucleus to demonstrate volumetric differences between normal aging and AD. As shown in Figure 2, controls had the highest bilateral caudate nucleus volume (mean: 3,777, SD = 1,017 mm³), followed by MCI (mean: 3,648, SD = 1,062 mm³; 3.42% lower), with AD subjects showing the lowest volumes (mean: 3,543, SD = 1,302 mm³; 6.20% lower); this confirmed the expected trend of AD < MCI < controls.

![Fig. 2. Volumetric Analysis (N = 400). Mean volumes for left, right, and pooled (left plus right) caudate volumes in the three diagnostic groups; error bars denote standard errors of the mean. Compared with controls caudate volume is lower bilaterally in MCI and still lower in AD. Also, right caudate volume is greater than left in the whole sample (*p* < 0.001), MCI (*p* = 0.005), and controls (*p* < 0.001), but the asymmetry is not detected in AD. Between groups, the right caudate was smaller in AD (−300 mm³, −8.47%, *p* = 0.001) and MCI (−163 mm³, −4.43%, *p* = 0.002) than in controls.](image-url)

We also calculated Cohen’s *d* statistics to tabulate the effect sizes for group differences in pooled (left plus right) volumes of the caudate nucleus, hippocampus (Morra et al., 2009), and lateral ventricles (Chou et al., 2009, 2010). For the continuous variables (MMSE, CDR, CDR-SB, and BMI scores), partial correlations (*r*-values) are reported from a multiple regression model that controlled for age and sex. *Significant results are highlighted with an asterisk (*p* < 0.05). The ventricular effects have an opposite sign to the others as the ventricles tend to expand in AD while other subcortical structures become smaller. As expected, effect sizes are generally largest for the hippocampus, then the lateral ventricles, then the caudate.

<table>
<thead>
<tr>
<th>Effect size comparisons</th>
<th>Caudate</th>
<th>Hippocampus</th>
<th>Ventricles</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohen’s <em>d</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD vs. Controls</td>
<td>0.378*</td>
<td>0.952*</td>
<td>−0.619*</td>
<td>193</td>
</tr>
<tr>
<td>MCI vs. Controls</td>
<td>0.252*</td>
<td>0.473*</td>
<td>−0.205*</td>
<td>286</td>
</tr>
<tr>
<td>AD vs. MCI</td>
<td>0.148</td>
<td>0.485*</td>
<td>−0.401*</td>
<td>279</td>
</tr>
<tr>
<td><em>r</em>-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>0.175*</td>
<td>0.349*</td>
<td>−0.205*</td>
<td>400</td>
</tr>
<tr>
<td>CDR</td>
<td>−0.062</td>
<td>−0.192*</td>
<td>0.143*</td>
<td>400</td>
</tr>
<tr>
<td>CDR-SB</td>
<td>−0.209*</td>
<td>−0.365*</td>
<td>0.225*</td>
<td>400</td>
</tr>
<tr>
<td>BMI</td>
<td>−0.177*</td>
<td>−0.197*</td>
<td>−0.028</td>
<td>373</td>
</tr>
</tbody>
</table>

Effect sizes for caudate nucleus, hippocampus, and lateral ventricle volumes. Cohen’s *d* statistics are shown to indicate effect sizes for group differences in the pooled (left plus right) volumes of the caudate nucleus, hippocampus (Morra et al., 2009), and lateral ventricles (Chou et al., 2009, 2010). For the continuous variables (MMSE, CDR, CDR-SB, and BMI scores), partial correlations (*r*-values) are reported from a multiple regression model that controlled for age and sex. *Significant results are highlighted with an asterisk (*p* < 0.05). The ventricular effects have an opposite sign to the others as the ventricles tend to expand in AD while other subcortical structures become smaller. As expected, effect sizes are generally largest for the hippocampus, then the lateral ventricles, then the caudate.

For continuous variables (MMSE, CDR, CDR-SB, and BMI scores), we computed a partial correlation from a multiple regression model, controlling for sex and age. The results are shown in Table 2.

A significant rightward asymmetry was found (the right caudate was larger) in controls (*p* < 0.001, asymmetry = 3.86%) and in MCI (*p* = 0.01, asymmetry = 2.13%). This asymmetry was no longer detected in AD. This rightward asymmetry is subtle and not always detected in MRI studies (see Discussion). Between groups, right caudate volume was smaller in AD (−300 mm³, −8.47%, *p* = 0.001) and MCI (−163 mm³, −4.43%, *p* = 0.002) versus controls, but

---

Table 2

<table>
<thead>
<tr>
<th>Effect size comparisons</th>
<th>Caudate</th>
<th>Hippocampus</th>
<th>Ventricles</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohen’s <em>d</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD vs. Controls</td>
<td>0.378*</td>
<td>0.952*</td>
<td>−0.619*</td>
<td>193</td>
</tr>
<tr>
<td>MCI vs. Controls</td>
<td>0.252*</td>
<td>0.473*</td>
<td>−0.205*</td>
<td>286</td>
</tr>
<tr>
<td>AD vs. MCI</td>
<td>0.148</td>
<td>0.485*</td>
<td>−0.401*</td>
<td>279</td>
</tr>
<tr>
<td><em>r</em>-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>0.175*</td>
<td>0.349*</td>
<td>−0.205*</td>
<td>400</td>
</tr>
<tr>
<td>CDR</td>
<td>−0.062</td>
<td>−0.192*</td>
<td>0.143*</td>
<td>400</td>
</tr>
<tr>
<td>CDR-SB</td>
<td>−0.209*</td>
<td>−0.365*</td>
<td>0.225*</td>
<td>400</td>
</tr>
<tr>
<td>BMI</td>
<td>−0.177*</td>
<td>−0.197*</td>
<td>−0.028</td>
<td>373</td>
</tr>
</tbody>
</table>

Effect sizes for caudate nucleus, hippocampus, and lateral ventricle volumes. Cohen’s *d* statistics are shown to indicate effect sizes for group differences in the pooled (left plus right) volumes of the caudate nucleus, hippocampus (Morra et al., 2009), and lateral ventricles (Chou et al., 2009, 2010). For the continuous variables (MMSE, CDR, CDR-SB, and BMI scores), partial correlations (*r*-values) are reported from a multiple regression model that controlled for age and sex. *Significant results are highlighted with an asterisk (*p* < 0.05). The ventricular effects have an opposite sign to the others as the ventricles tend to expand in AD while other subcortical structures become smaller. As expected, effect sizes are generally largest for the hippocampus, then the lateral ventricles, then the caudate.
left caudate volume was not significantly different between diagnostic groups.

We compared MCI subjects who converted to AD with those who remained in the MCI diagnostic group at future follow-ups. At both 1- and 2-year time points, we found significantly greater baseline atrophy in the right caudate among MCI-to-AD converters compared with nonconverting MCI subjects (p = 0.033, 0.99%; p = 0.032, 2.41% respectively for 1- and 2-year follow ups).

Next we examined correlations between caudate nucleus volume and a variety of clinical measures including MMSE, CDR, and Logical Memory scores, with the goal of ranking their strength of association. Poorer MMSE scores at baseline (p < 0.001, N = 400) and changes in MMSE scores from baseline at a 1-year follow-up (p = 0.008, N = 400) were both significantly correlated with greater right caudate atrophy at baseline, when all diagnostic groups were pooled. Similarly, greater atrophy of the right caudate was also correlated with poorer global CDR scores (p = 0.013, N = 400). Both right and left caudate atrophy were associated with worse “sum of boxes” CDR scores (p = 0.010, for the left; p > 0.001, for the right, N = 400) and with poorer Delayed Logical Memory recall (p = 0.019, for the left; p = 0.005, for the right, N = 395). Poorer Immediate Logical Memory recall was also correlated with right caudate atrophy (p = 0.004, N = 395).

### Group difference maps

Next we created 3D maps showing the significance levels (p-values) for local caudate atrophy in AD and MCI versus controls (Fig. 3), after controlling for age and sex. Separate maps are shown for the left and right caudate nucleus.

Relative to controls, greater atrophy was detected in AD (left: p = 0.022, corrected; right: p = 0.001, corrected; N = 200) and MCI (left: p = 0.023, corrected; right: p = 0.012, corrected; N = 300), but not in AD versus MCI (left: p = 0.383, corrected; right: p = 0.083, corrected; N = 300).

### Baseline association maps of clinical scores

Next we created maps of correlations for the clinical and cognitive scores that showed significant associations in our volumetric analyses (Fig. 3). These map-based results were largely in line with the analysis of overall volumes. Tests where poorer scores were associated with volume reduction...
bilaterally (CDR-SB, Delayed Logical Memory) also showed a significant bilateral effect in the maps. Bilateral atrophy was associated with greater impairment measured by CDR (right: \( p = 0.029 \), corrected; \( N = 400 \), CDR-SB (left: \( p = 0.002 \), corrected; right: \( p < 0.001 \), corrected; \( N = 400 \)), and in Delayed Logical Memory (left: \( p = 0.015 \), corrected; right: \( p = 0.003 \), corrected; \( N = 395 \)). Poorer Immediate Logical Memory scores (right: \( p = 0.006 \), corrected; \( N = 395 \)) were associated with right caudate atrophy only, as in the volumetric analysis.

### 2.4. Associations with clinical scores 1 year later

To investigate the predictive value of our caudate maps, we associated baseline caudate anatomy with change in MMSE and CDR scores between baseline and 1-year later (Fig. 4). As suggested by our volumetric analysis, right caudate atrophy was highly associated with baseline MMSE scores in the full sample (\( p = 0.001 \), corrected, \( N = 400 \)) and with 1-year change in MMSE scores in the MCI group (left: \( p = 0.037 \); right: \( p = 0.001 \), corrected, \( N = 200 \)). Associations of 1-year change in MMSE scores were not significant for either AD or control groups; however we cannot rule out that these associations may be found over a longer time interval, greater than 1 year. Although associations with baseline CDR and CDR-SB were significant (Fig. 3), associations with 1-year change in these scores were not significant after multiple comparisons correction.

### 2.5. Greater atrophy in MCI-to-AD converters

To investigate the predictive value of local caudate nucleus atrophy, we created significance maps showing the association between caudate nucleus atrophy at baseline among MCI subjects who converted to AD (converters) versus those who remained classified as MCI (nonconverters) 1 or 2 years after their baseline scan (Fig. 4). Right caudate volumes were lower at baseline for those who had converted to AD 1 year later (\( p = 0.008 \), corrected; \( N = 174 \)) or 2 years later (\( p = 0.035 \), corrected; \( N = 50 \)).

### 2.6. Associations with CSF biomarkers

Right but not left caudate atrophy (Fig. 4) was associated with CSF tau concentrations (right: \( p = 0.021 \), corrected; \( N = 216 \), AD: \( N = 59 \), MCI: \( N = 102 \), controls: \( N = 55 \)). Additionally, correlation analyses within each diagnostic group revealed that both left and right caudate nucleus volume are correlated with CSF tau concentration in the AD subjects (left: \( r = 0.239 \), \( p = 0.034 \); right: \( r = 0.370 \), \( p = 0.002 \); \( N = 59 \)) but not in the MCI or control groups (\( r < 0.1 \), \( p > 0.05 \)). Correlation maps of associations between radial atrophy and, concentrations of-tau and amyloid beta and ratio measures of tau/amyloid beta and p-tau/amyloid beta were also assessed, but these correlation maps did not pass permutation-based corrections for multiple comparisons. It would be interesting to see if the CSF associations

---

**Fig. 4.** Maps of associations with MMSE scores at baseline and 1 year later, MCI-to-AD conversion, and CSF concentrations of tau. 3D maps show areas of significant associations between local volumetric atrophy in the caudate and MMSE scores at baseline and after a 1-year follow-up interval, with \( p \)-values color-coded at each surface voxel. Associations were only detected in the right caudate nucleus for baseline MMSE scores (Right: \( p = 0.001 \), corrected). Change in MMSE score after a 1-year follow-up interval was bilaterally associated in the MCI group only (Left: \( p = 0.037 \); Right: \( p = 0.001 \), corrected; \( N = 200 \)). Those who converted from MCI to AD after 1 year and 2 years showed greater right caudate atrophy at baseline than those who did not (1-year conversation: \( p = 0.008 \), corrected; 2-year conversation: \( p = 0.035 \), corrected). For CSF levels of tau at time of scanning, associations were significant for the right but not left caudate for the tau CSF biomarker (\( p = 0.036 \), corrected, \( N = 216 \)).
hold up on a group-by-group level since other ADNI laboratories have found that tau concentrations increase progressively from controls to MCI to AD (Shaw et al., 2009); however, our current study did not have enough power to detect associations within each diagnostic group separately (as sample sizes with available CSF data are only around 50 subjects per group).

2.7. Associations with additional clinical variables

Significant associations were also found in maps for age (left: \( p = 0.039 \), corrected; right: \( p = 0.001 \), corrected; \( N = 400 \)) and sex (left: \( p < 0.001 \), corrected; right, \( p < 0.001 \), corrected; \( N = 400 \)), with smaller caudates in men than women (Fig. 5), and smaller caudates in older than younger people. Examining emotional domains, we also found significant associations between more impaired Geriatric Depression scores and atrophy of the right caudate nucleus (right, \( p = 0.004 \), corrected; \( N = 395 \)). In our analysis of motor function, the presence of abnormal gait was significantly associated with volume reductions in the right caudate nucleus (right, \( p = 0.027 \), corrected; \( N = 395 \); however, maps of associations between atrophy and the presence of tremor in our subjects were not significant.

In line with recent tensor-based morphometry findings that higher BMI is associated with brain atrophy in the elderly, caudate atrophy was found here to be associated with higher BMI in our full sample (left: \( p < 0.001 \), corrected; right: \( p = 0.002 \), corrected; \( N = 393 \)) and in the AD subjects considered separately (left: \( p = 0.015 \), corrected; right: \( p = 0.005 \), corrected; \( N = 96 \)). Within-group associations for MCI and control groups were not significant.

Maps of associations with the FTO genotype, which is linked to higher BMI and with brain atrophy in elderly people (Ho et al., 2010b), were not significant after permutation-based correction for multiple comparisons. Maps associating local caudate nucleus volume with APOE genotype were also not significant after multiple comparisons correction.

2.8. Ranking effect sizes for different covariates

Results from our permutation-corrected significance maps agreed with our volumetric results for each statistical test performed, while providing a more detailed spatial profile of regional atrophy. Adjusting for age and sex in these regressions did not affect these associations materially, but slightly boosted effect sizes (data not shown). To rank the covariates in terms of their effect sizes, we ran the maps with randomly selected samples of 400, 300, 200, 96, and 50 subjects, to find the reduced sample size needed to detect each effect, using the results of the permutation test to correct for multiple comparisons. A sample size of 96 was chosen instead of 100 to allow inclusion of the associations for BMI in the AD group only, which consisted of 96 subjects. Only 50 subjects were needed to detect the BMI

![Fig. 5. Associations between caudate atrophy and additional clinical variables.](image-url)
effect in the AD group (left: \(p = 0.024\), corrected; right: \(p = 0.022\), corrected; \(N = 50\)). In a representative sample, 96 subjects were needed to detect effects of BMI (left: \(p = 0.0031\), corrected; right: \(p = 0.006\), corrected; \(N = 96\)), baseline CDR-SB scores (right: \(p = 0.037\), corrected; \(N = 96\)), and the comparison of AD versus controls (right: \(p = 0.038\), corrected; \(N = 96\)). Effects of baseline MMSE scores were first detected at 200 subjects (right: \(p < 0.001\), corrected; \(N = 200\)).

3. Discussion

This study demonstrates that 3D maps of the caudate nucleus localize statistically significant regions of volume reduction associated with (1) diagnosis and clinical and cognitive scores, (2) future decline in MMSE scores at 1-year follow-up in the MCI group, (3) conversion from MCI to AD 1 and 2 years after scanning, (4) baseline CSF concentrations of tau, and (5) additional clinical measures including BMI, Geriatric Depression, and abnormal gait. The variables most strongly associated with caudate atrophy were found using a reduced N analysis, namely (in this rank order): BMI in the AD group only, BMI in all subjects, CDR-SB scores, group difference between AD and controls, and MMSE scores at baseline. It is interesting to note that cardiovascular measures and CDR-SB scores, which rely on interviews from both patients and informants, were more powerfully associated with caudate volume reductions in our 3D maps, compared with seemingly more direct measures of AD such as the MMSE. Although some of these measures are highly correlated and may be viewed as redundant, we felt that it was important to take a broad approach initially, in analyzing multiple measures of cognition and health. This has practical significance for the planning of clinical trials, which currently use a variety of similar tests, and also allows direct comparison with other studies that each may use different tests.

Caudate atrophy was generally localized to anterior regions and was specific to the right caudate for several variables assessed. Caudate atrophy in AD (vs. controls) was 2-fold greater on the right than the left. Further study is needed to confirm the greater vulnerability of the right caudate. The natural asymmetry of the caudate (right 3.9% larger than the left in controls) has been found in many but not all studies, and large samples are needed to detect it. Most literature supports a rightward (right larger than left) asymmetry in caudate nucleus volume (Giedd et al., 1996; Ifthikharuddin et al., 2000; Larisch et al., 1998; Peterson et al., 1993; Yamashita et al., 2009), but some studies have found leftward, gender-specific, or no asymmetry (Gunning-Dixon et al., 1998; Szabo et al., 2003). Based on our current findings and previous literature, it may be possible to view asymmetry findings in brain structures as a general indicator of poor brain health. Psychosis has been linked to genes involved in the development of cerebral asymmetry (Laval et al., 2004) and differences in laterality have been found to covary with, or predict, individual differences in susceptibility to stress pathology and drug sensitivity (Carlson and Glick, 1989). The presence or change in direction of asymmetry has been reported (but not always consistently) in disorders ranging from schizophrenia (Crow et al., 1996), ADHD (Castellanos et al., 1996), bipolar (Javadapour et al., 2010) affective mood disorders (Baumann et al., 1999), and Tourette syndrome (Klieger et al., 1997).

There may be some skepticism that the caudate is a natural place to look for clinical correlations in AD, but it may give additional power to detect associations with cognition and future decline, if used along with structures that are more commonly studied. For example, it is widely accepted that the hippocampus — target of most published AD morphometry studies — is challenging to segment reliably, with a notorious 2-fold difference in mean hippocampal volumes across studies and imaging centers (Frisoni et al., 2010). Prior studies using the same radial distance mapping technique with other brain structures in ADNI have also found correlations between hippocampal atrophy (Morra et al., 2009), ventricular expansion (Chou et al., 2009, 2010) and clinical variables such as MMSE, CDR, and CDR-SB scores. There are also differences between diagnostic groups in the volumes of all of these structures (AD vs. controls, MCI vs. controls). As shown in Table 2, and as might be expected, hippocampal volumes gave the greatest effect sizes for detecting group differences within ranges supported by previous studies (Shen et al., 2010), and for detecting correlations with clinical scores. Effect sizes for lateral ventricle volumes also gave relatively strong effect sizes, although they were generally weaker than the hippocampal effect sizes. Caudate effect sizes, while significant, were generally lower than those for the ventricles and hippocampus. As one exception to this rank order, the caudate gave marginally better effect sizes than ventricular volume for distinguishing MCI from controls. Even so, we did not perform statistical tests for differences in effect sizes across structures, as a stringent multiple comparisons correction would be needed. Despite these generally weaker effects, the current study of the caudate also found significant associations between reduced volume and MCI-to-AD conversion, and with decline in MMSE scores after a 1-year follow-up in the MCI group, whereas these associations were not detectable in studies of ventricular expansion (Chou et al., 2009) or hippocampal atrophy (Morra et al., 2009a, 2009b), using the same ADNI subject pool (with 240 and 490 subjects, respectively); the latter study also used the same Adaboost machine learning algorithm, applied to hippocampal segmentation. This is an intriguing finding, suggesting caudate volume at baseline may predict later clinical changes relatively well, perhaps because caudate atrophy is only severe when the disease is rapidly progressing.
The caudate has previously been implicated in AD through multiple lines of evidence including histological findings of tau and amyloid accumulations (Braak and Braak, 1990). Volumetric studies find reductions in caudate nucleus volume with normal aging (Jernigan et al., 2001; Krishnan et al., 1990; Raz et al., 2003), in AD versus controls (Rombouts et al., 2000), and in many other neurological, psychiatric and neurodevelopmental conditions including autism (Turner et al., 2006), fragile X syndrome (Gothelf et al., 2008), elderly depression (Butters et al., 2008), and HIV (Becker et al., 2006). Caudate lesions can produce deficits in executive control and cognitive processing speed (Rubin, 1999). Interestingly, a recent study also found subcortical brain volume reductions, encompassing the caudate nucleus, in elderly individuals with a higher BMI (Raji et al., 2009).

The anterior caudate head is the most closely involved in traditionally studied cognitive processes and integrates major inputs from the dorsolateral and orbitofrontal cortices, involved in attention and planning (Cummings, 1995). Disturbances in the hemispheric asymmetry of the caudate head are associated with ADHD and schizophrenia (Ballmaier et al., 2008; Blanton et al., 1999; Hynd et al., 1993). One study found significant caudate volume reductions in depressed elderly and created 3D maps using the same technique here, except on manually derived segmentations; however, their maps were not significant after permutation-based correction for multiple comparisons (Butters et al., 2008).

Beyond the caudate and the medial temporal lobes, additional basal ganglia and subcortical gray matter structures are implicated in AD and age-related neurodegeneration. Putamen and thalamus volumes are reduced in AD, and volume decreases correlate with cognitive deficits (de Jong et al., 2008). Putamen volumes are reduced in frontotemporal lobe degeneration, but not in AD (Looi et al., 2009). A recent study of individuals with Down’s syndrome—a group at heightened risk for developing AD—found significant associations in those who developed AD and bilateral volume reductions in the hippocampus, caudate, right amygdala, and right putamen (Beacher et al., 2009). In vivo imaging with [11C]-PIB PET, which is thought to map the profile of amyloid deposition in the living brain, has correlated increased amyloid signal with increased atrophy in the hippocampus and amygdala (Frissoni et al., 2009). In addition, post-mortem studies have localized diffuse amyloid plaques in the caudate nucleus in AD (Ikonomovic et al., 2008).

In a PET study of dopamine D1 and D2 receptors, researchers found a 14% reduction in mean signal for D1 receptors for the caudate and putamen in AD, but no difference in D2 receptor signal (Kemppainen et al., 2000). The basal ganglia also have the highest iron content of any brain region, which is interesting because AD patients exhibit disrupted iron metabolism that may relate to oxidative damage and higher iron levels are found in the caudate and putamen in AD patients compared with controls (Bartzokis et al., 2000).

A common limiting factor in neuroimaging studies of AD is the labor-intensive task of manually segmenting subcortical structures on large MRI datasets, requiring long-term effort from expertly trained neuroanatomists. Fortunately, automated methods now exist to efficiently segment many subcortical structures including the hippocampus and caudate nucleus (Fischl et al., 2002; Powell et al., 2008; Yushkevich et al., 2006). Here we used an automated method based on adaptive boosting, that we recently developed and validated (Morra et al., 2008). Using this method, we efficiently created caudate nucleus segmentations in 400 ADNI brain MRI scans. The resulting 3D maps had enough statistical power to detect local volume differences associated with a wide variety of clinical variables in AD. Our maps detected associations between caudate atrophy and CSF levels of tau, which have not always been associated with other MRI-derived measures in the same cohort (Chou et al., 2009). CSF levels of tau are significantly associated with higher rates of temporal lobe atrophy (Leow et al., 2009), but only showed a trend level of association with the rate of hippocampal atrophy (Schuff et al., 2009). In prior amyloid PET studies, amyloid deposition has been noted in the caudate nucleus (Ikonomovic et al., 2008). Here we did not detect a relation between the level of CSF amyloid and the volume of the caudate. In the future, a direct PET-based brain measure of amyloid in the caudate may correlate better with atrophy than CSF-based measures of amyloid burden. Limited sample sizes may have also made it hard to detect an association; even so, the sample was large enough for tau levels to be reliably correlated with maps of atrophy for the right caudate nucleus in the pooled sample and bilaterally for the volumetric summaries in the AD group.

A possible confounding factor in some voxel-based morphometric studies of the caudate is the potential for misregistration of anatomy across subjects along the ventricles, especially in elderly subjects with substantial brain degeneration. Because the caudate nucleus is a gray matter structure that lies just below the boundary of the lateral ventricles, some volumetric studies may be confounded by poor or biased registration along the lateral ventricular surface. This effect is unlikely in the current study as our measures of radial atrophy are intrinsic, meaning they are independent of whether the structure is translated in space. They are computed from a centerline traced down the center of the structure, and do not rely on the correct registration of images across subjects, as some voxel-based averaging methods do.

The etiology of caudate atrophy in AD and MCI is of interest. One probable explanation is that we are assessing neuronal atrophy, secondary to the accumulation of amyloid plaque pathology and tau neurofibrillary tangles (Braak and Braak, 1990). The caudate may also experience a loss of afferent projections from other brain regions. As such, a...
limitation of the study is that the associations between caudate atrophy and cognition may in reality be mediated by atrophy occurring in other brain structures, such as the hippocampus. The intent in this study was to identify correlates of caudate atrophy, but it is likely that the more immediate causes of the clinical differences may be atrophy (or functional compromise) of structures elsewhere in the brain, not necessarily those assessed here.

A further limitation of the current study is that the available sample sizes made it necessary to use a pooled sample including both patients and controls for several analyses. Pooling subjects allowed us to achieve sufficient statistical power to detect important findings, but there is also a risk of recovering group sampling effects due to the differences between diagnostic groups in cognitive measures. Where sample sizes and effect sizes were sufficient for within-group correlations to be detected, we did find statistically significant associations in 1-year change in MMSE score in the MCI group and with BMI in the AD group. It is possible that there are additional within-group associations for these and other measures that did not survive the sample size reduction. It can also be argued that AD, MCI, and healthy elderly controls represent stages on a continuum — in that case, binning the subjects into diagnostic groups could limit the power of correlations found across the spectrum and lead to false negative results.

In this study, we report correlations between atrophy and cognitive or CSF-derived measures in a pooled ADNI sample (combining patients with AD, MCI and controls), while also reporting several within-groups correlations split by diagnosis (“disaggregated” analyses of AD subjects only, for example). Both types of analysis are complementary and each has its own limitations. When analyzing a mixed cohort of subjects, our goal is to determine (1) the cognitive correlates of atrophy in the entire sample, and (2) whether the chosen biomarker of disease burden is linked with decline across the full spectrum of controls, MCI, and AD subjects. As the whole cohort is arguably a continuum, it is vital to look beyond diagnostic categories to analyze relationships between atrophy and particular measures of cognitive function. A within-group analysis of AD subjects only, for example, may not be able to detect these correlations due to a “restricted range” effect seen when looking at a limited range of differences in brain structure and cognitive performance. By running split analyses only, many important correlations may be missed. For instance, brain atrophy correlates well with CSF-derived measures of tau pathology across the continuum from healthy aging to MCI and to AD. However, it is possible that no correlation will be detected if we subselect a group such as MCI, with a very narrow range of disease burden. Furthermore, it is a fallacy to preselect groups that are defined partly on cognitive measures, and then later test for correlations with these same cognitive scales. If the selection criterion for the group correlates with the variable of interest, we will find many false negatives due to the truncated range, making the results uninterpretable.

Pooled analyses have limitations to consider, as well. When cognitive measures are correlated with diagnosis, correlations in a pooled cohort may show similar patterns to those found when directly comparing AD subjects with controls. Another consideration when using pooled cohorts such as ADNI, is that correlations may be influenced by the proportion of subjects in each diagnostic group. By design, ADNI oversampled subjects with cognitive decline, including a 1:2:1 ratio of AD: MCI: controls. Correlations in this particular sample may not be detected to the same degree in other population studies with different proportions of subjects or within each diagnostic group. For these reasons, we chose to use a combination of both pooled and split analyses, which have complementary value in understanding the cognitive and pathological correlates of atrophy.

Because it can be difficult to tease apart contributions of each individual brain region in a disease such as AD, studying effects of the disease on specific anatomical structures, beyond those most commonly targeted, may help to piece together relevant circuitry affected by AD pathology. Multiple AD biomarkers can be used in a mutually reinforcing way to identify subjects most likely to decline to AD in a sample (Kohannim et al., 2010), with caudate atrophy may inform the classification of AD, MCI or the prediction of future decline. Another line of work involves genome-wide association studies of caudate volume in large subject cohorts such as ADNI. This information will improve our understanding of the mechanistic processes involved in brain aging and AD.

Disclosure statement

The authors have no potential financial or personal conflicts of interest including relationships with other people or organizations within 3 years of beginning the work submitted that could inappropriately influence this work.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health, Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly, and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer, Inc., F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., and Wyeth, as well as nonprofit partners the Alzheimer’s Association and Alzheimer’s Drug Discovery Foundation, with participation from the US Food
References


Gunter, J., Bernstein, M., Botorwski, B., Felmlee, J., Bledez, D., Mallozi, R., 2006. Validation testing of the MRI calibration phantom for the Alzheimer’s Disease Neuroimaging Initiative study. ISMRM 14th Scientific Meeting and Exhibition, Seattle, WA.


