Comparing Hippocampal Atrophy in Alzheimer’s Dementia and Dementia with Lewy Bodies

Nicole Chow\textsuperscript{a,b} Dag Aarsland\textsuperscript{e,h,k} Hedieh Honarpisheh\textsuperscript{d} Mona K. Beyer\textsuperscript{i}
Johanne H. Somme\textsuperscript{j} David Elashoff\textsuperscript{c} Arvid Rongve\textsuperscript{f} Ole B. Tysnes\textsuperscript{g}
Paul M. Thompson\textsuperscript{a,b} Liana G. Apostolova\textsuperscript{a,b}

\textsuperscript{a}Department of Neurology, \textsuperscript{b}Imaging Genetics Center, Laboratory of Neuroimaging, David Geffen School of Medicine and \textsuperscript{c}Division of General Internal Medicine and Health Services Research, UCLA, Los Angeles, Calif., and \textsuperscript{d}Pathology and Laboratory Medicine, Yale School of Medicine, New Haven, Conn., USA; \textsuperscript{e}Institute of Clinical Medicine, University of Oslo, Oslo; \textsuperscript{f}Department of Psychiatry, Haugesund Hospital, Haugesund; \textsuperscript{g}Department of Neurology, Haukeland University Hospital, Bergen; \textsuperscript{h}Centre for Age-Related Diseases, and Norwegian Centre for Movement Disorders, Stavanger University Hospital, Stavanger, Norway; \textsuperscript{i}Department of Neurology, Cruces University Hospital, Baracaldo, Spain; \textsuperscript{j}Department of Neurobiology, Ward and Care Services, Alzheimer’s Disease Research Centre, Karolinska Institutet, Stockholm, Sweden

Key Words
Alzheimer’s disease · Dementia with Lewy bodies · Hippocampus · Magnetic resonance imaging · Atrophy

Abstract

\textbf{Background/Aims:} Dementia with Lewy bodies (DLB) and Alzheimer’s disease (AD) are the two most common neurodegenerative dementias. During the early stages, clinical distinction between them is often challenging. Our objective was to compare hippocampal atrophy patterns in mild AD and mild DLB. We hypothesized that DLB subjects have milder hippocampal atrophy relative to AD subjects.

\textbf{Methods:} We analyzed the T1-weighted magnetic resonance imaging data from 113 subjects: 55 AD, 16 DLB and 42 cognitively normal elderly (normal controls, NC). Using the hippocampal radial distance technique and multiple linear regression, we analyzed the effect of clinical diagnosis on hippocampal radial distance, while adjusting for sex and age. Three-dimensional statistical maps were adjusted for multiple comparisons using permutation-based statistics with a threshold of $p < 0.01$.

\textbf{Results:} Compared to NC, AD exhibited significantly greater atrophy in the cornu ammonis (CA)1, CA2–3 and subicular regions bilaterally while DLB showed left-predominant atrophy in the CA1 region and subiculum. Compared directly, AD and DLB did not reveal statistically significant differences. \textbf{Conclusion:} Hippocampal atrophy in mildly impaired DLB subjects is less severe than atrophy seen in mildly impaired AD subjects. Both AD and DLB show predominant atrophy of the CA1 subfield and subiculum.

Introduction

Dementia with Lewy bodies (DLB) – the second most common neurodegenerative dementia after Alzheimer’s disease (AD), has a complex clinical presentation including psychiatric, motor, sleep and autonomic disturbances in addition to cognitive impairment [1]. Prognosis and quality of life in DLB are generally even poorer than in...
AD [2], as are health-related costs [3] and mortality [4]. DLB patients have a different response to drug treatment and are at a particularly high risk of developing severe hypersensitivity reactions to neuroleptic drugs [5].

Differentiating DLB from other dementias, including AD, is of particular clinical importance. However, this distinction can be difficult, especially early in the disease course when the classic DLB clinical profile may not yet be fully developed [6]. Consequently, many pathologically confirmed cases of DLB have been clinically misdiagnosed as AD premortem [7–10]. While the clinical consensus criteria for DLB were recently revised [1] and preliminary evidence suggests that the newly revised criteria have greater sensitivity [11], a systematic evaluation of their sensitivity and specificity is not yet available.

The ability of a range of biomarkers to aid in the differential diagnosis of DLB is actively being explored. The most established method to date, dopamine transporter single-photon-emission computerized tomography [12], is relatively expensive and is not readily available at all centers. Structural neuroimaging using magnetic resonance imaging (MRI) is among the most established biomarkers for AD and is now used for the diagnosis of prodromal AD [13]. Novel sensitive MRI analytic techniques have been recently developed and have enabled us to identify presymptomatic and early symptomatic structural changes in AD [14–16].

Few studies have compared the MRI changes in DLB and AD; most have used either visual rating, the region-of-interest approach, or whole brain imaging analysis methods such as voxel-based morphometry or cortical thickness approaches to measure cortical or subcortical atrophy in DLB compared to normal controls (NC) or AD. These types of studies find that relative to AD, DLB subjects showed significantly less atrophy in the orbitofrontal and temporal lobes [17–19]. In addition, temporal lobe and hippocampal atrophy are less pronounced in DLB than in AD [19–22], but the magnitude of the differences is small, precluding the ability of standard volumetric hippocampal assessments to readily distinguish between AD and DLB. Yet one recent study suggested that visual scoring of medial temporal atrophy on MRI obtained approximately 1.5 years prior to death might provide discriminatory power for distinguishing patients with moderately severe pathologically confirmed AD (Mini-Mental State Examination, MMSE 13.8 ± 4.54) from those with moderately severe pathologically confirmed DLB (MMSE 13.3 ± 7.83) and mild vascular cognitive impairment (MMSE 22.8 ± 4.36) [21]. Another study applied a recently developed hippocampal partial subfield segmentation technique to the 3-tesla MRI data of 16 AD (mean MMSE 21.5, range 16–27) and 16 DLB (mean MMSE 18, range 15–27) patients and 16 NC subjects (mean MMSE 29, range 26–30) [23]. Their measurements were limited to the three most anterior contiguous coronal slices from the hippocampal body. Despite significant technical difficulties in ascertaining the CA1 subfield boundary, the authors reported significantly smaller CA1 area in AD when compared to DLB.

Most prior MRI studies comparing DLB and AD have included patients with moderately severe dementia. However, recent advances in biomarker development and structural neuroimaging, in particular, now allow us the unprecedented opportunity to detect structural changes in the presymptomatic stages for other dementing disorders such as AD [15, 24]. In this study, we analyzed the imaging data of DLB and AD subjects in the mild dementia stages. Our objective was to compare hippocampal atrophy patterns in mild AD and mild DLB. We hypothesized that DLB subjects would have milder hippocampal atrophy relative to subjects diagnosed with AD, with potentially greater involvement of the CA2–3 subfields, as was recently suggested by one postmortem study [21].

Methods

Subjects

The Dementia study in western Norway (DemWest) is a multisite longitudinal study of the natural history and biological correlates of dementia. Details of inclusion and assessment procedures have been published previously [25]. In brief, 196 subjects with mild dementia – defined as an MMSE [26] score of 20 or higher – were recruited from referrals to all geriatric-medicine, old-age-psychiatry and neurology outpatient clinics in Rogaland and Hordaland counties in western Norway between March 2005 and March 2007. The study was later enriched with more DLB subjects, including some in the moderate dementia stages (MMSE range 18–20). Standardized clinical instruments were employed to detect and rate cognition, psychiatric and motor symptoms as previously described [25]. Diagnosis of AD was made according to the National Institute of Neurological and Communication Disorders and Stroke and the AD and Related Disorders Association (NINCDS-ADRDA) criteria [27] and diagnosis of DLB according to the revised DLB consensus criteria [1]. Two research psychiatrists independently applied the diagnostic criteria twice – at baseline and after 1 year. In cases of disagreement, and whenever more than one set of operationalized diagnostic criteria was met, final diagnostic ascertainment was made based on consensus between the 2 physicians after careful review of all available information. All DemWest patients were actively recruited...
to autopsy. To date, 7 patients have undergone postmortem exams using standard methods and diagnostic criteria, as previously described [28, 29]. In all 7 cases, the pathological diagnosis agreed with the premortem clinical diagnosis. Our study analyzed the imaging data of 55 AD and 16 DLB DemWest subjects who provided a baseline structural MRI scan of sufficient quality for imaging analyses.

Our control group consisted of 42 cognitively normal elderly subjects who were enrolled in the ParkWest study – another longitudinal project in western Norway. These NC subjects were scanned during the same time period on the same scanners and with the same imaging protocol as our DemWest subjects. ParkWest NC were free from parkinsonism, dementia, major depression and psychosis and scored within the cognitively normal range on the detailed ParkWest neuropsychological battery which has been previously described [30]. Both studies were approved by the regional Norwegian committee for medical research ethics for western Norway. All subjects gave written informed consent for the participation in the study after procedures had been explained in detail in accordance with the Declaration of Helsinki.

Imaging Data Collection and Analysis

Subjects were scanned at 5 different sites located in Stavanger, Haugesund, Haraldsplass, Bergen and Arendal. The following protocols were used:

- Stavanger: 1.5 T Philips Intera (Best, The Netherlands), repetition time (TR)/echo time (TE) 10.0/4.6 ms, flip angle 30°, 2-mm slices, 1-mm gap, number of excitations (NEX) 2, matrix 256 × 256, nominal resolution is 1 × 1 × 1.28 mm.
- Haugesund: 1.5 T Philips Intera (Best, The Netherlands), TR/TE 20.0/4.6 ms, flip angle 30°, 1-mm slice thickness with no gap, NEX 1, matrix 256 × 256, nominal resolution is 1 × 1 × 1.46 mm.
- Haraldsplass: 1.5 T General Electric Signa Excite (Milwaukee, Wisc., USA), TR/TE 8.2/3.1 ms, flip angle 7°, 1-mm slice thickness, 1-mm gap, NEX 1, matrix 256 × 256, nominal resolution is 1 × 1 × 1.29 mm.
- Bergen: 1.5 T General Electric Signa Excite (Erlangen, Germany), TR/TE 8.2/3.1 ms, flip angle 7 degrees, 1-mm slice thickness with no gap, NEX 1, matrix 256 × 256, nominal resolution is 1 × 1 × 1.33 mm.
- Arendal: 1.0 T Philips Intera (Best, The Netherlands), TR/TE 25/6.9 ms, flip angle 30°, 2-mm slice thickness with no gap, NEX 1, matrix 256 × 256, nominal resolution is 1 × 1 × 1.28 mm.

T2-weighted and fluid-attenuated inverted recovery (FLAIR) sequences were collected to evaluate subjects for strokes and/or structural lesions. Subjects with these findings were excluded from our imaging analyses, as were those with baseline scan artifacts or scans of insufficient quality. Our final cohort of DemWest subjects consisted of 71 patients. There were no significant differences in age, sex, education, disease duration and MMSE scores between DemWest subjects who underwent an MRI and those who did not.

Individual MRI scans were automatically registered to the International Consortium for Brain Mapping 53 (ICBM53) template, an average of 53 normal adult brains, using a 9-parameter transformation. This step orient each brain volume into the ICBM53 standardized coordinate system by rotating and globally scaling to correct for differences in head tilt and head size between subjects. Next, hippocampi were manually traced on coronal slices by one researcher (HH, intrarater reliability Cronbach’s alpha = 0.9) blinded to the age, sex, education, MMSE score and diagnosis of subjects, and following our detailed hippocampal tracing protocol as previously described [31]. The traces included the hippocampus proper, dentate gyrus and subiculum. Traces were converted into hippocampal contours and transformed into 3-dimensional parametric surface mesh models, which were then separated into top and bottom components [32]. These mesh models assured normalization of the spatial frequency of the digitized surface points. Next, a medial core, threading down the center of the hippocampus, was computed. Radial distance was measured from the medial core to the surface of the hippocampus. Each radial distance value was recorded at the corresponding surface coordinate point. These resulting individual hippocampal radial distance maps were combined across subjects to create group average distance maps for quantitative comparisons of surface morphology between diagnostic groups [32].

Intra- and Inter-Scanner Reliability Analyses

Human phantom scanning of 3 cognitively normal individuals was performed to test scanner reliability. Each volunteer was scanned twice with all scanners included in this study. We performed interscanner reliability analyses. The hippocampi were manually traced (JHS, intrarater reliability Cronbach’s alpha = 0.896) and the volumes obtained as previously described. Intersite Cronbach’s alpha for hippocampal volumes was 0.967.

Statistical Methods

One-way analyses of variance (ANOVA) with post hoc Bonferroni correction for multiple comparisons were run to examine between-group diagnostic differences in continuous variables such as age, education and MMSE scores. A χ2 test was used to determine differences in sex distribution. The effect of diagnosis on hippocampal radial distance was studied by means of linear regression while correcting for demographic variables that showed significant between-group differences. Our 3-dimensional statistical maps were adjusted for multiple comparisons, using permutation-based statistics with a threshold of p < 0.01.

Results

Mean demographic data for the diagnostic groups NC (n = 42), AD (n = 55) and DLB (n = 16) are shown in Table 1. The AD and DLB subjects were well matched for overall cognitive impairment as assessed by MMSE (AD 23.9 ± 2.2, DLB 23 ± 2.7, p = 0.4), the Clinical Dementia Rating (CDR) global score (AD 0.76 ± 0.30, DLB 0.80 ± 0.25, p = 0.671) and CDR Sum of Boxes (AD 4.39 ±
Hippocampal Atrophy in AD and DLB

1.93, DLB 5.03 ± 1.99, p = 0.260). Significant differences in age were found, with the DLB group being the oldest and the NC group being the youngest (p < 0.0001). Sex distribution was also significantly different with the DLB group having more men than women, while the opposite was true for the AD and NC groups (p = 0.026). As expected, NC subjects had a significantly higher mean MMSE score (28.81 ± 1.1) than AD (23.9 ± 2.2) and DLB subjects (23.0 ± 2.7) (p < 0.0001), but there was no significant difference between AD and DLB. Age and sex were included as covariates in our radial distance multiple regression models.

Compared to NC, the AD group exhibited significantly greater atrophy in the CA1, CA2–3 and subicular region bilaterally (left $p_{corrected} < 0.0001$, right $p_{corrected} = 0.0001$; fig. 1, 2nd row), while the DLB group showed greater left than right atrophy localizing mainly to the CA1 region and subiculum (left $p_{corrected} = 0.0004$, right $p_{corrected} = 0.056$; fig. 1, 3rd row). Quantitatively, both DLB and AD groups had between 10–40% smaller radial distance in the statistically significant areas relative to NC. The AD versus DLB comparison did not reveal statistically significant between-group differences (fig. 1, bottom row). The average and standard deviation radial distance maps for each group can be seen in figure 2.

Discussion

Differentiating DLB from AD early in the disease course can be challenging, as the complete clinical profile of DLB may not yet be fully developed. Here, we used an advanced surface-based technique to determine if characteristic structural changes are present in the mild dementia stages of DLB and AD that might aid clinicians in their differential diagnosis. We observed significant hippocampal atrophy in both diagnostic groups compared to NC. Our mild AD subjects showed greater atrophy in all hippocampal subfields as previously demonstrated [33]. Our mild DLB subjects showed significant atrophy of the CA1 and subiculum relative to NC. Another study, by Sabatelli et al. [22], using the same hippocampal radial distance technique, investigated the atrophy pattern in 14 mild-to-moderate DLB subjects (MMSE range 13–29) relative to NC and reported atrophy restricted to the anterior portions of CA1. Despite enrolling a cognitively milder DLB cohort (MMSE range 18–26), our study was able to detect more widespread atrophy of the CA1 and subiculum.

Our DLB and AD groups were well matched for overall cognitive impairment, allowing a fair diagnostic comparison. However, with the sample size of the DLB group (n = 16), we failed to show statistically significant differences between the groups. To further analyze the failure to detect significant differences between AD and DLB, we constructed a confidence interval (CI) of hippocampal volumes to determine the range of likely values for which we would expect to find a difference between the groups. The difference between the left volumes of AD and DLB had a 95% CI of 59.0 ± 340.2 mm³ and the right volume difference had a 95% CI of –69.2 ± 349.4 mm³. These intervals indicate that, at most, the difference between groups was 10%. Using the average volume for each subject, the CI for the difference in means is –5.1 ± 315 mm³. With this interval, the difference between the groups was calculated to be 7.5%, which is in agreement with the percent AD vs. DLB difference map seen in the last row of Figure 1. Therefore we can conclude that in our study, the AD and DLB groups are indeed very similar in terms of hippocampal volume. The calculated CIs exclude the possibility that these groups are substantially different from one another. While these findings seem to disagree with those reported in other studies [23], there

<table>
<thead>
<tr>
<th>Variable</th>
<th>NC (n = 42)</th>
<th>AD (n = 55)</th>
<th>DLB (n = 16)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>69.86±6.756 (50–82)</td>
<td>74.76±7.371 (55–89)</td>
<td>78.13±6.479 (69–88)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Sex, M:F</td>
<td>14:28</td>
<td>13:42</td>
<td>9:7</td>
<td>0.446</td>
</tr>
<tr>
<td>Education, years</td>
<td>10.21±2.526 (7–17)</td>
<td>9.48±2.543 (7–18)</td>
<td>8.53±1.803 (7–12)</td>
<td>0.053</td>
</tr>
<tr>
<td>MMSEa</td>
<td>28.81±1.110 (27–30)</td>
<td>23.85±2.155 (19–29)</td>
<td>23.00±2.683 (18–26)</td>
<td>&lt;0.0001 (AD vs. DLB, p = 0.4)</td>
</tr>
<tr>
<td>CDRb</td>
<td>N/A</td>
<td>0.76±0.302 (1–2)</td>
<td>0.80±0.254 (1–1)</td>
<td>0.671</td>
</tr>
<tr>
<td>CDR Sum of Boxesb</td>
<td>N/A</td>
<td>4.39±1.931 (2–9)</td>
<td>5.03±1.986 (3–10)</td>
<td>0.260</td>
</tr>
</tbody>
</table>

a Two subjects (one AD and one DLB) had MMSE = 19 and one DLB subject had MMSE = 18.

b CDR and CDR Sum of Boxes available for 15 DLB subjects.
are several plausible explanations for this. Firstly, DLB is a heterogeneous disorder with the majority of cases showing, in addition to Lewy bodies, pathological characteristics of AD such as amyloid plaques and neurofibrillary tangles. Mixed DLB/AD cases frequently show hippocampal atrophy and neurofibrillary tangle pathology of a severity similar to that seen in AD [21]. As such, studies that have relatively small sample sizes could have a study population skewed towards either pure or mixed DLB. Neither our study nor the study by Firbank et al. [23] had postmortem diagnostic verification to ascertain the pathological diagnoses of our subjects. The study by Firbank et al. [23] used a recently developed method for subfield tracing that involves CA1, CA2 and CA3/4 subfield differentiation on only 3 contiguous coronal sections from the hippocampal structure, with tracing beginning on the first slice where the hippocampal head is no longer visible. Such sparse sampling of the hippocampal structure potentially driven by the tediousness and substantial technical difficulties behind subfield
tracing, even at 4-tesla [23, 34, 35], may or may not generalize well to the whole subfield or to the whole hippocampus. Upon close inspection of the AD versus DLB comparison (fig. 1, bottom row), one can appreciate that there are significant between-group differences in the suprolateral CA1 area adjacent to the right hippocampal head. This is likely very similar to the area where Firbank et al. [23] reported greater atrophy in AD versus DLB. As in AD [37], neurofibrillary tangles have been linked with hippocampal atrophy in DLB [21]. Yet, postmortem subjects with DLB show only mild neuronal loss in CA1 and minimal neurofibrillary tangle and neurite pathology in the CA2–3 regions [36].

Several strengths and limitations of our study should be recognized. Major strengths of the study include its design, the well-characterized patient cohort using standardized instruments to validate core and suggestive DLB features, and its focus on the mild dementia stages. The state-of-the-art imaging analysis is another strength; it allows the identification of focal, regionally specific, disease-associated differences. One limitation of our study is the lack of pathological validation for all cases. However, DemWest subjects are routinely approached for postmortem diagnostic assessment. Postmortem diagnostic confirmation was completed on the first 7 cases that came to autopsy, and in all cases, the clinical and pathological diagnosis were in agreement. Finally, we should also recognize that despite our best efforts to control for between-group variability by entering age and sex as potential confounders in the linear regression analyses, some residual variance due to these and other demographic imbalances might still be present.

**Acknowledgment**

We would like to acknowledge the Research Council of Norway (grant number 177966), the Western Norway Regional Health Authority (grant number 911218), the Norwegian Parkinson’s Disease Association and research grants from Stavanger University Hospital and the Western Norway Regional Health Authority (grant number 911464) for financial support in the design and conducting of the study as well as the collection and management of obtained data. Our work is also supported by NIH (EB008281, AG020098, RC2 AG036535) and NIA P50 A16570, which provided funding for the analysis and interpretation of data. All sources of support were involved in the preparation, review and approval of the manuscript. Finally, we would like to acknowledge Ketil Oppeled and Turi Olene Dalaker for their contribution to the human phantom scans.

**Fig. 2.** Average radial distance maps and standard deviation maps of each diagnostic group.
References


