Neurobiology of Aging 49 (2017) 204–213

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Medial temporal lobe subregional morphometry using high resolution MRI in Alzheimer's disease

David A. Wolk [a,b,*], Sandhitsu R. Das [a], Susanne G. Mueller [c], Michael W. Weiner [c], Paul A. Yushkevich [d], for the Alzheimer's Disease Neuroimaging Initiative [1]

[a] Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA
[b] Penn Memory Center, University of Pennsylvania, Philadelphia, PA, USA
[c] Department of Veterans Affairs Medical Center, Center for Imaging of Neurodegenerative Diseases, San Francisco, CA, USA
[d] Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA

ARTICLE INFO

Article history:
Received 11 April 2016
Received in revised form 18 September 2016
Accepted 18 September 2016

Keywords:
Preclinical Alzheimer disease
Prodromal Alzheimer disease
Subfields
Perirhinal cortex
MRI
Biomarkers

ABSTRACT

Autopsy studies of Alzheimer's disease (AD) have found that neurofibrillary tangle (NFT) pathology of the medial temporal lobe (MTL) demonstrates selective topography with relatively stereotyped subregional involvement at early disease stages, prompting interest in more granular measurement of these structures with in vivo magnetic resonance imaging. We applied a novel, automated method for measurement of hippocampal subfields and extrahippocampal MTL cortical regions. The cohort included cognitively normal (CN) adults (n = 86), early mild cognitive impairment (n = 43), late MCI (n = 22), and mild AD (n = 40) patients from the Alzheimer's Disease Neuroimaging Initiative (ADNI). For pseudolongitudinal analysis of the continuum from preclinical to mild AD dementia, the groups were further divided according to amyloid status based on positron emission tomography. Specific subregions associated with the early NFT pathology of AD were more sensitive to preclinical and early prodromal AD than whole hippocampal volume while more diffuse involvement was found in later stages. In particular, BA35, the first region associated with NFT deposition, was the only region to discriminate preclinical AD from amyloid negative cognitively normal adults ("normal aging"). In general, patterns of atrophy in the pseudolongitudinal analysis largely recapitulated Braak staging of NFTs within the MTL.

1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative condition in which impairments of episodic memory are an early hallmark symptom. Memory loss is likely driven, in part, by the well-described early neurofibrillary tangle (NFT) pathology in the medial temporal lobe (MTL), the central anatomic structure supporting episodic memory (Arnold et al., 1991; Braak and Braak, 1991; Squire, 2004). Although measurements of the hippocampus have proven to be one of the most reliable biomarkers of AD, both for prediction of cognitive decline in mild cognitive impairment (MCI) and for tracking disease progression, more granular measurement of MTL substructures may prove to enhance sensitivity at different disease stages (Pluta et al., 2012).

Indeed, in early stages, NFT pathology does not uniformly involve the MTL and hippocampus (Arnold et al., 1991; Braak and Braak, 1991). The first region thought to display this pathology is perirhinal cortex (PRC), particularly Brodmann Area 35 (BA35), which Braak and Braak referred to as transentorhinal cortex (Braak and Braak, 1991). This is then followed by involvement of the entorhinal cortex (ERC) and then cornu ammonis 1 (CA1) subfield of the hippocampus proper. As such, more global measures of MTL structure may obscure these more specific regional changes at early stages.

With the goal of developing more sensitive measures of early AD pathology, a number of in vivo approaches for measurement of hippocampal subfields and cortical MTL structures have been

* Corresponding author at: Penn Memory Center, 3615 Chestnut Street, #212A, Philadelphia, PA 19104, USA. Tel.: 215 662-7810; fax: 215 662-7812.
E-mail address: david.wolk@uphs.upenn.edu (D.A. Wolk).

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

0197-4580/$ – see front matter © 2016 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neurobiolaging.2016.09.011

developed using magnetic resonance imaging (MRI) (Apostolova et al., 2006; Iglesias et al., 2015; Kerchner et al., 2010; Mueller and Weiner, 2009; Van Leemput et al., 2009; Wisse et al., 2012; Yushkevich et al., 2015; Zeineh et al., 2003). However, high field strength (7T MRI), relatively long acquisition times (~30 minutes), and labor intensive manual annotation limit the practical utility of some of these methods for larger cohorts and clinical populations. Use of a high in-plane resolution (~0.4 × 0.4 × 2 mm^3), but anisotropic, T2-weighted MRI sequence orthogonal to the long axis of the hippocampus, has proven a useful compromise, as these scans can be obtained in approximately 8 minutes at 3T field strength. These scanning parameters allow for visualization of the internal structures of the hippocampus, not visible on a typical research-quality T1-weighted scan (1 × 1 × 1 mm^3).

Although a number of manual protocols for segmentation of the above T2-weighted sequence have been published (La Joie et al., 2013; Mueller et al., 2007, 2008; Shing et al., 2011; Yushkevich et al., 2015), our group developed an automated technique for labeling hippocampal subfields and cortical MTL regions [(Pluta et al., 2012; Yushkevich et al., 2010, 2015); see (Iglesias et al., 2015) for a different automated approach]. Using this method, measurement of selected subregions, including CA1 and BA35, has demonstrated greater sensitivity to MCI status than whole hippocampal volumes (Pluta et al., 2012; Yushkevich et al., 2015).

The present article describes the relative involvement of different MTL subregions across a range of cognitive impairment associated with AD. In addition to assessing the relative sensitivity of these measures in clinically classified groups, we explored whether regional atrophy patterns correspond to distribution of NFT deposition described in autopsy studies. To do so, we examined differences in regional MTL atrophy across the continuum from cognitively normal (CN) without evidence of AD pathology, or "amyloid negative" based on florbetapir-PET, to preclinical AD (CN adults with evidence of cerebral amyloid, or "amyloid positive"), prodromal AD ("amyloid positive" MCI patients), and early AD dementia. The current contribution represents the most comprehensive analysis of morphometric data measured from an MTL-dedicated sequence obtained in a multisite manner as part of ADNI.

2. Materials and methods

2.1. Participants

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies and nonprofit organizations, as a $60 million, 5-year public–private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as 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 is the result of efforts of many coinvestigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the United States and Canada. The initial goal of ADNI was to recruit 800 subjects, but ADNI has been followed by ADNI-GO and ADNI-2. To date these 3 protocols have recruited over 1500 adults, aged 55–90, to participate in the research, consisting of CN older individuals, people with early or late MCI, and people with early AD. The follow-up duration of each group is specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

For this study, subjects were selected from the subset of the ADNI 2 participants in which high-resolution T2-MRI was collected at the 20 sites using scanners developed by Siemens Medical Systems. This acquisition was developed after initiation of ADNI 2 data collection as a subproject with the goal of comparing the efficacy and added value of MTL subregional measurements relative to whole hippocampal and other standard volumetric measurements (Iglesias et al., 2015; Mueller et al., 2013). Clinical categorization of study participants into CN, early MCI (EMCI), late MCI (LMCI), and AD were as described in detail at www.adni-info.org and used standard clinical criteria for MCI and AD (Landau et al., 2012; McKhann et al., 1984; Petersen, 2004). MCI patients were all amnestic, either single domain or multi-domain. Division into early and late MCI groups was based solely on education-adjusted scores for the delayed paragraph recall subscore of the Wechsler Memory Scale-Revised Logical Memory II. Thus, EMCI patients straddled the boundary between normal memory and LMCI.

In light of the delayed initiation of the high-resolution T2-MRI substudy, many of the included MRI's did not correspond to baseline clinical assessments (n = 149 of 191). In these instances, diagnosis and psychometric testing closest to the time of the scan was used. As MCI patients were not systematically divided into EMCI and LMCI at these follow-up visits, we used the same Wechsler Memory Scale-Revised Logical Memory II cutoff to dichotomize into these 2 categories. Further, the control group included subjects designated as significant memory concern, a category added after the initiation of ADNI 2 recruitment based on the presence of subjective memory complaints and a Cognitive Change Index score ≥16 (Saykin et al., 2006), but psychometric testing within the normal range. As we did not observe clear differences in control participants with or without such subjective complaints, these groups were collapsed into the CN group.

2.2. Psychometric testing

The following measures were included in the current analysis: Wechsler Memory Scale-Revised Logical Memory I and II (Wechsler, 1987), Rey Auditory Verbal Learning Test (Rey, 1964), the Trail Making Test (Trails A and Trails B) (Reitan, 1958), category fluency (Animals) (Butters et al., 1987), and Boston Naming Test (Kaplan et al., 1983). Auditory verbal learning test measures include 5 immediate memory trials and 5- and 30-minute delayed recall trials and a 30-minute delayed recognition measure. In addition, the Clinical Dementia Rating scale (Morris, 1993) was obtained in all participants. Demographic and clinical characteristics of the subject pool are described in Table 1.

2.3. MRI image acquisition

All subjects were scanned with a T1-MRI protocol optimized for best contrast to noise in a feasible acquisition time (Jack et al., 2008; Leow et al., 2006). Raw data had an acquisition matrix of 192 × 192 × 166 and voxel size 1.25 × 1.25 × 1.2 mm^3. Zero-filled reconstruction (i.e., sinc interpolation) resulted in a 256 × 256

Table 1
Group demographic, psychometric, and biomarker data

Demographics	CN (n = 86)	EMCI (n = 43)	LMCI (n = 22)	AD (n = 40)
Age	74.3 (6.9)	74.4 (7.1)	74.1 (7.4)	74.3 (8.7)
Education	16.9 (2.2)	16.7 (2.9)	16.7 (2.8)	16.6 (2.4)
LM delay	14.7 (3.3)	12.0 (3.5)***	4.3 (3.2)***	1.5 (2.9)***
Trails A (s)	32.0 (12.9)	36.9 (16.6)	38.4 (17.8)	55.4 (31.4)***
Trails B (s)	84.6 (62.1)	92.7 (48.7)	111.24 (62.6)	171.3 (90.3)***
BNT total	28.5 (2.0)	27.5 (3.4)	27.0 (3.9)	25.5 (6.0)***
Category fluency	21.4 (5.7)	18.7 (5.1)**	16.8 (5.0)**	12.2 (4.6)***
AVLT (trials 1–5)	47.8 (10.9)	43.0 (12.7)*	31.5 (9.3)***	23.4 (8.3)***
AVLT 5-min delay	9.8 (3.5)	7.8 (3.8)**	4.6 (3.1)***	1.4 (1.9)***
AVLT 30-min delay	8.2 (4.3)	6.1 (4.8)*	1.9 (2.7)***	0.5 (1.7)***
Florbetapir PET (SUVR)	1.13 (0.3)	1.16 (0.2)	1.29 (0.21)***	1.36 (0.21)***

All statistics are in comparison to CN group; *$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$. Note that 1 LMCI, and 3 AD patients did not complete Trails B and 1 AD did not complete the AVLT 30-min delayed recall.
Key: AD, Alzheimer's disease; AVLT, Auditory Verbal Learning Test; BNT, Boston Naming Test; CN, cognitively normal adults; EMCI, early mild cognitive impairment; LM, logical memory; LMCI, late mild cognitive impairment; PET, positron emission tomography.

matrix and voxel size of 0.9375 × 0.9375 × 1.2 mm^3. In addition, T2-MRI with high in-plane resolution in the oblique coronal orientation prescribed to be perpendicular to the long axis of the hippocampal formation was also acquired (Fig. 1). The MRI sequence parameters were repetition time/echo time 8020/50 ms, 0.4 × 0.4 × 2.0 mm^3 resolution, minimum 24 slices, acquisition time: 8.1 minutes. Further details on ADNI imaging protocols can be found at http://adni.loni.usc.edu/methods/documents/mri-protocols/.

2.4. Image processing

Image quality of high-resolution T2-MRI was assessed by a manual rater using the following criteria: (1) Each image was assigned a quality rating between 0 and 5 based on the severity of noise, motion, and ringing artifacts. Images with a rating >2 were included in the study; and (2) Images where part of the posterior hippocampal complex was not included because of a field of view that was too far anterior were excluded. For each participant, data from the earliest available MRI session that met the selection criteria above were used. High-resolution T2-MRI was labeled using the Automatic Segmentation of Hippocampus Subfields (ASHS) software package (Yushkevich et al., 2010, 2015). Briefly, this method uses a multiatlas label fusion technique in combination with a learning-based error correction module to produce a fully automated segmentation of hippocampal subfields along the entire length of the hippocampal formation, as well as segmentation of some MTL cortical regions. Candidate segmentations of a subject's MRI are obtained using high-dimensional mapping to multiple manually labeled atlas images. These are then fused into a consensus segmentation, taking into account the degree of similarity between the subject image and atlas images. Patterns of systematic segmentation errors introduced in this procedure are learned a priori using training data and are corrected in a further postprocessing step, to generate the final segmentation.

Cornu Ammonis (CA1/2/3) and dentate gyrus (DG) subfields of the hippocampal formation, subiculum, and entorhinal (ERC), parahippocampal (PHC), and perirhinal (PRC) cortices (BA35 and BA36) were labeled in each subject. The segmentation protocol used for atlas set generation and reliability of automated labeling for these subregions was previously reported (Yushkevich et al., 2015). Further details of the protocol can be found in supplementary material for this prior report: http://onlinelibrary.wiley.com/store/10.1002/hbm.22627/asset/supinfo/hbm22627-sup-0001-suppinfo.docx?v=1&s=7ce9ee4821c571dcbf6ad9981cc9682c41350fac. An example of the automatic segmentation is shown is Fig. 1.

Voxelwise cortical thickness maps for the entire brain were generated from T1-MRI using a diffeomorphic registration-based thickness estimation method described in (Das et al., 2009), as implemented in the ANTS Cortical Thickness pipeline, evaluated in Tustison et al. (2014). Briefly, a one-to-one mapping between the white–gray matter interface and the gray matter–cerebrospinal fluid interface is derived, and thickness is defined as the distance between corresponding points. T1 and T2 images were registered with FLIRT software (Jenkinson and Smith, 2001), as part of the hippocampal subfield segmentation pipeline (ASHS) (Yushkevich et al., 2015). We visually verified that there were no registration failures. Thickness maps were then integrated over the cortical labels obtained from T2-MRI to generate average cortical thickness in the ERC and PRC subregions. Whole hippocampal volumes were generated from the anatomical image of each subject using the multi-atlas label fusion technique previously described (Wang and Yushkevich, 2013).

2.5. Florbetapir-PET

Florbetapir-PET scanning consisted of four 5-minute dynamic frames, acquired 50–70 minutes after injection and is described in detail on-line (adni.loni.ucla.edu/about-data-samples/image-data/). Processed Florbetapir PET data were made publicly available by ADNI and used the following automated processing pipeline (adni.loni.ucla.edu/research/pet-post-processing/). Briefly, after coregistration with T1-weighted MRI, an SUVR was calculated by taking the mean uptake over voxels in 4 regions of interest (ROIs), frontal, anterior/posterior cingulate, lateral parietal, and lateral temporal cortices, and then dividing these regions by mean uptake over voxels of the entire cerebellum. A composite measure taking the mean SUVR of the 4 cortical ROIs was used to determine whether a participant was "amyloid positive" or "amyloid negative." We used a previously established cutoff, 1.11, for this measure (Landau et al., 2012). SUVR values of this composite measure are included in Table 1.

2.6. Statistical analysis

Morphometric measurements from MTL subregions were entered into statistical analysis using IBM SPSS 22.0. For hippocampal subfields, volumes produced by ASHS were used. Because of the uncertainty associated with the lateral boundaries of MTL cortical structures, cortical thickness was used for ERC, PRC, and PHC regions instead of volumes. To reduce the number of comparisons, mean values of right and left hemispheres were calculated for each MTL region. In addition, given poorer reliability of CA2 and CA3, based on our prior work (Yushkevich et al., 2015), a combined CA label was used. Consistent with the fact that CA1 accounts for most of this volume, results were highly similar whether we used CA1 or the entire CA region.

The discriminative ability of the imaging biomarkers derived from high-resolution T2-MRI was examined by comparing volumes or cortical thickness between CN group and each of the 3 patient groups: EMCI, LMCI, and AD. Further comparisons were made across the spectrum of AD by comparing amyloid negative CN adults with amyloid positive CN participants (i.e., preclinical AD), EMCI (i.e., early prodromal AD), LMCI (i.e., late prodromal AD), and

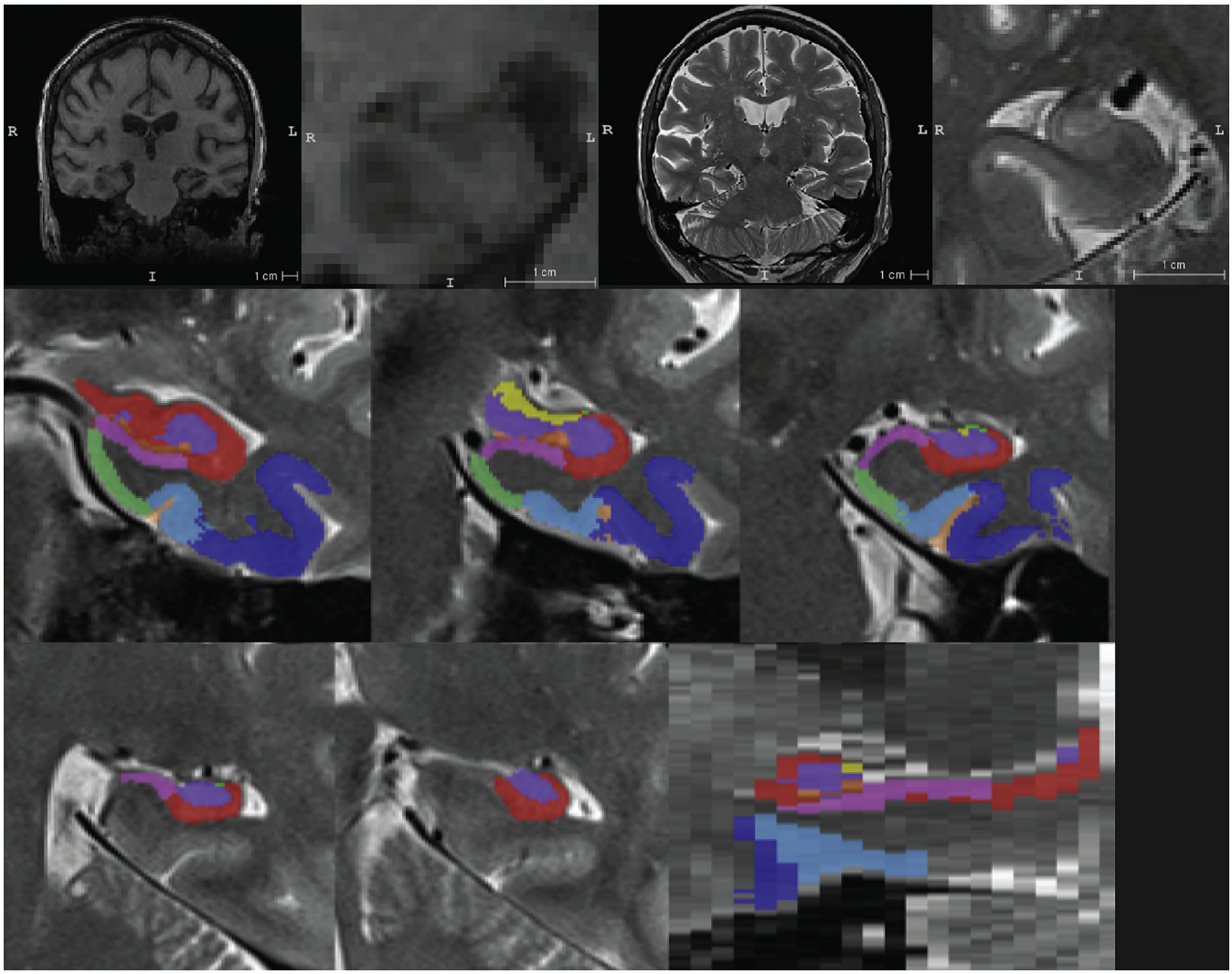

Fig. 1. Top row from left to right: T1-MRI, T1-MRI zoomed into hippocampus in coronal view, T2-MRI, and T2-MRI zoomed into hippocampus in oblique coronal view. Middle and bottom row show an example of automatic segmentation of MTL subregions. Abbreviations: MTL, medial temporal lobe; MRI, magnetic resonance imaging.

AD (i.e., mild AD dementia). All values were adjusted by age and education based on linear regression in amyloid negative CN adults. For volumes, intracranial volume produced by ASHS was also used to adjust values. Direct group comparisons used 1-sided *t* tests.

3. Results

3.1. Clinical group discrimination

Results of the comparisons of symptomatic patients versus CN adults are tabulated in Table 2. Average hippocampal subfield volumes as well as thickness in MTL cortex were progressively smaller in EMCI, LMCI, and AD relative to controls. Importantly, we found that subregion-specific measurements were sensitive to the earliest symptomatic disease stage, EMCI, but not whole hippocampal volume. PRC thickness, particularly that of BA36 ($t[70.1] = 4.0$, $p < 0.001$)], had a strong group effect when compared with controls. All other extrahippocampal MTL regions were significantly different as well, but with smaller effect sizes (ERC: $t[127] = 3.0$, $p < 0.01$; BA35: $t[63.6] = 2.4$, $p < 0.05$; PHC: $t[127] = 2.0$, $p < 0.05$). Hippocampal regions were generally more marginally significant (subiculum: $t[127] = 2.3$, $p < 0.05$; CA: $t[62.6] = 1.7$, $p < 0.05$; and DG: $t[61.5] = 1.8$, $p < 0.05$). Notably, whole hippocampal volume was not significantly different between these groups ($t[59.9] < 1$, $p > 0.1$)].

In LMCI, although both BA35 ($t[106] = 5.3$, $p < 0.001$) and BA36 ($t[106] = 4.1$, $p < 0.001$) continued to be quite sensitive to clinical status, CA ($t[106] = 7.2$, $p < 0.001$) displayed the largest effect size. Indeed, the pattern of atrophy appeared to be much more diffuse at this stage, with DG ($t[106] = 5.7$, $p < 0.001$), SUB ($t[106] = 5.2$, $p < 0.001$), and whole hippocampal volume ($t[24.6] = 4.6$, $p < 0.001$) demonstrating significant group discrimination. Similarly, at the AD stage, most regions displayed highly significant group differences when compared to controls with CA having the largest effect size, but only marginally relative to whole hippocampal volume.

3.2. Groups dichotomized by amyloid status

Our primary analysis was to examine the relative effect on MTL subregional structure of AD along the continuum of preclinical disease to mild dementia. To so, a "normal aging" group (amyloid negative CN adults) was compared to preclinical AD (amyloid positive CN adults), early and late prodromal AD (amyloid positive EMCI and LMCI, respectively), and mild AD dementia (amyloid positive AD) based on amyloid status using florbetapir PET

Table 2
Comparisons to control group after adjustment for age and education, as well as ICV for volume measurements

Region	CN (n = 86)	EMCI (n = 43)	LMCI (n = 22)	AD (n = 40)
CA				
Volume (SD) mm^3	1323.5 (147.4)	1261.9 (213.6)	1059.0 (182.0)	1027.8 (190.4)
T value		1.7	7.2	8.7
p-value		<0.05	<0.001	<0.001
Dentate gyrus				
Volume (SD) mm^3	786.8 (86.9)	747.8 (129.4)	654.9 (127.4)	628.0 (116.5)
t value		1.8	5.7	8.5
p-value		<0.05	<0.001	<0.001
Subiculum				
Volume (SD) mm^3	440.0 (61.2)	413.7 (62.5)	365.3 (53.0)	344.4 (51.6)
t value		2.3	5.2	8.6
p-value		<0.05	<0.001	<0.001
ERC				
Thickness (SD) mm	3.25 (0.42)	2.99 (0.54)	2.78 (0.56)	2.68 (0.44)
t value		3.1	4.3	7.0
p-value		<0.01	<0.001	<0.001
BA35				
Thickness (SD) mm	2.58 (0.45)	2.32 (0.63)	1.99 (0.53)	2.02 (0.55)
t value		2.4	5.3	5.6
p-value		<0.05	<0.001	<0.001
BA36				
Thickness (SD) mm	3.35 (0.41)	2.99 (0.51)	2.91 (0.55)	2.78 (0.56)
t value		4.0	4.1	5.7
p-value		$2e^{-5}$	$4e^{-5}$	$2e^{-7}$
PHC				
Thickness (SD) mm	2.25 (0.45)	2.06 (0.61)	1.84 (0.40)	1.96 (0.53)
t value		2.0	3.8	3.1
p-value		<0.05	<0.001	<0.01
Hippo				
Volume (SD) mm^3	2144.6 (218.8)	2074.3 (364.8)	1758.9 (381.3)	1691.0 (307.6)
t value		1.2	4.6	8.4
p-value		>0.1	<0.001	<0.001

Key: AD, Alzheimer's disease; CN, cognitively normal adults; EMCI, early mild cognitive impairment; ERC, entorhinal cortex; ICV, intracranial volume; LMCI, late mild cognitive impairment; PHC, parahippocampal; SD, standard deviation.

(Table 3). We only included florbetapir scans that were obtained within 1-year of the MRI scan (mean: 135.6 ± 157.1 days) used for MTL subregion analysis, and, thus, some individuals in the prior analysis were excluded. In light of the 8 ROIs examined per group comparison, statistical effects that would survive Bonferroni correction ($p < 0.006$) were noted with bold font in Table 3.

The only region to reach significance in the preclinical AD group relative to "normal aging" was BA35 ($t[62] = 2.0, p < 0.05$). No other region approached significance. Interestingly, in absolute terms measured by percent difference, BA35 (8.1%) and PHC (5.2%) displayed the greatest atrophy in the preclinical AD (all other regions <1%). Indeed, several regions, including CA and whole hippocampus volume was actually somewhat larger in the preclinical AD group.

In contrast, in the early prodromal AD stage, most regions displayed some degree of decreased volume or thickness relative to the "normal aging" group. All extrahippocampal MTL regions reached significance (BA35: $t[21.0] = 2.9$, $p < 0.01$; BA36: $t[59] = 3.4$, $p < 0.001$; ERC: $t[59] = 2.9$, $p < 0.01$; PHC: $t[59] = 3.5$, $p < 0.001$). Although all of the hippocampal subregions were also significant, these effects were more marginal (subiculum: $t[59] = 1.8$, $p < 0.05$; CA: $t[20.0] = 1.8$, $p < 0.05$; DG: $t[20.0] = 2.0$, $p < 0.05$). Notably, whole hippocampal volume did not significantly differ with the "normal aging" group but did display a nonsignificant trend ($t[18.4] = 1.5$, $p = 0.07$). Extrahippocampal MTL regions also displayed generally greater percent differences with the "normal aging" group than hippocampal structures (~12–18% vs. ~6–8%).

In late prodromal AD and mild AD dementia, there was much more diffuse involvement of MTL structures, and all regions reached statistical significance in comparison with "normal aging". Indeed, CA was the region of largest effect size in the late prodromal AD comparison ($t[53] = 6.7$, $p < 0.001$), and both CA and whole hippocampus best discriminated groups in mild AD dementia (t's = 7.0, p's < 0.001). Nonetheless, even in the latter group, BA35 displayed the greatest percent difference with the "normal aging" group. Fig. 2 plots the evolution of atrophy in select regions across disease stage with this pseudolongitudinal data. Note that adjustment for gender had no substantive impact on these comparisons.

We also calculated the estimated sample size required to discriminate the amyloid positive groups from the "normal aging" group with a significance criterion of 0.05 (1-sided) and a power of 0.8 to correctly reject the null hypothesis if it is false. We assumed equal samples sizes from both groups (see Table 3). As suggested by the above statistical comparisons, extrahippocampal MTL regions would allow for smaller sample sizes in preclinical and early prodromal AD than hippocampal subfields and whole hippocampal volume although this advantage would be lost during the late prodromal and mild dementia stages. Importantly, BA35 appeared to allow for detection of disease effects with a reasonable number of participants in these very early stages (i.e., n = 86 in preclinical AD and n = 34 in prodromal AD).

Finally, we examined whether there were statistically significant differences in discrimination between the subfields with largest effect sizes versus whole hippocampus by comparing areas under the curve (AUC). We did find that the BA35 had a significantly higher AUC than whole hippocampus in discriminating "normal aging" from preclinical AD (0.65 vs. 0.44, $p < 0.05$). In early prodromal AD, PHC displayed a trend toward significantly greater discrimination (0.76 vs. 0.60, $p = 0.08$). In late prodromal AD, AUC's favored BA35 (0.93) and CA (0.93) relative to whole hippocampus (0.89), but these differences did not reach significance ($p > 0.1$).

4. Discussion

There are 2 primary findings from the current work. The first is that more granular measurement of MTL subregions provides enhanced sensitivity to early stages of AD compared with whole hippocampal measurements. The second is that the pattern of atrophy within the MTL associated with different degrees of AD severity, from preclinical to dementia, generally recapitulates the topographic pattern of NFT deposition reported in the histology literature. Indeed, regions associated with the earliest NFT pathology, particularly PRC and ERC, displayed significant atrophy in preclinical and early prodromal AD (Braak and Braak, 1991). Taken together, these data hold promise for the value of MTL subregional measurements as a biomarker sensitive to the earliest neurodegenerative effects of disease that could serve an important role in tracking disease progression in preclinical and prodromal AD intervention studies. As this represents the first experience with our fully automated segmentation pipeline in a multi-site context, the present findings also support the practicality of this approach in future clinical studies.

4.1. Selective involvement of MTL subregions in preclinical and prodromal AD

A primary motivation for measurement of subregions within the MTL is the potential increased sensitivity to early AD relative to

Table 3
Comparison to CN, amyloid negative group; all values adjusted for age and education and volumes adjusted for ICV

Region	CN (n = 44), amyloid neg	CN (n = 20), amyloid pos	EMCI (n = 17), amyloid pos	LMCI (n = 11), amyloid pos	AD (n = 26), amyloid pos
CA					
Volume (SD) mm^3	1311.8 (125.9)	1340.2 (176.9)	1208.7 (224.3)	1018.3 (148.2)	1069.2 (160.6)
% Difference from CN		−2.2	7.9	22.4	18.5
t value		<1	1.8	6.7	7.0
p-value		>0.1	<0.05	**<0.001**	**<0.001**
Sample size		N/A	80	8	12
Dentate gyrus					
Volume (SD) mm^3	784.7 (74.3)	780.6 (103.2)	715.8 (133.4)	634.6 (106.1)	661.3 (94.6)
% Difference from CN		0.5	8.7	19.2	15.7
t value		<1	2.0	5.5	6.1
p-value		>0.1	<0.05	**<0.001**	**<0.001**
Sample size		>1000	64	12	14
Subiculum					
Volume (SD) mm^3	437.8 (56.9)	442.1 (72.4)	407.7 (64.0)	357.8 (47.9)	360.5 (43.8)
% Difference from CN		−0.1	6.9	18.3	17.7
t value		<1.0	1.8	4.3	6.0
p-value		>0.1	<0.05	**<0.001**	**<0.001**
Sample size		N/A	102	14	14
ERC					
Thickness (SD) mm	3.23 (0.45)	3.27 (0.44)	2.84 (0.55)	2.89 (0.48)	2.71 (0.45)
% Difference from CN		−1.2	12.1	10.5	16.1
t value		<1	2.9	2.2	4.7
p-value		>0.1	**<0.01**	<0.05	**<0.001**
Sample size		N/A	42	46	20
BA35					
Thickness (SD) mm	2.70 (0.40)	2.48 (0.41)	2.22 (0.64)	1.86 (0.46)	2.00 (0.50)
% Difference from CN		8.1	17.8	31.1	26.0
t value		2.0	2.9	6.0	6.4
p-value		<0.05	<0.01	**<0.001**	**<0.001**
Sample size		86	34	10	12
BA36					
Thickness (SD) mm	3.35 (0.48)	3.34 (0.33)	2.92 (0.52)	2.94 (0.53)	2.66 (0.48)
% Difference from CN		0.0	12.8	12.2	20.6
t value		<1.0	3.4	2.8	6.4
p-value		>0.1	**<0.01**	<0.01	**<0.001**
Sample size		>1000	36	40	14
PHC					
Thickness (SD) mm	2.31 (0.43)	2.19 (0.42)	1.89 (0.43)	1.76 (0.34)	2.00 (0.58)
% Difference from CN		5.2	18.2	23.8	13.4
t value		<1	3.5	4.0	2.5
p-value		>0.1	**<0.01**	**<0.001**	<0.05
Sample size		312	28	14	70
Hippo					
Volume (SD) mm^3	2156.8 (176.2)	2193.5 (234.1)	2003.0 (402.1)	1768.2 (252.5)	1748.4 (266.3)
% Difference from CN		−1.7	7.1	17.9	18.9
t value		<1.0	1.5	6.0	7.0
p-value		>0.1	0.07	**<0.001**	**<0.001**
Sample size		N/A	104	10	10

p-values in bold reflect comparisons that survive Bonferroni correction ($p < 0.006$).
Key: AD, Alzheimer's disease; CN, cognitively normal adults; EMCI, early mild cognitive impairment; ERC, entorhinal cortex; ICV, intracranial volume; LMCI, late mild cognitive impairment; PHC, parahippocampal; SD, standard deviation.

whole hippocampal measurement (Fox et al., 1996; Jack et al., 1992, 2011). Indeed, several MTL subregions displayed stronger discrimination between preclinical and early symptomatic patients from CN adults, but caution must be applied as the power to detect statistical differences in discrimination was lacking once groups were divided based on amyloid status. Nonetheless, BA35 was the only region to discriminate preclinical AD from amyloid negative CN adults ("normal aging") although BA35/36, ERC, and PHC discriminated early prodromal AD from this group. In the latter comparison, hippocampal subfields displayed more modest discrimination and whole hippocampus trend level discrimination. By later clinical stages, this advantage was less evident. Such differences in sensitivity to these early disease changes have implications for size of cohorts that may need to be recruited in the context of an intervention study. This is demonstrated in the sample size calculations in which, for example, 17 individuals per group (34 total) would be needed for a reasonably powered study discriminating early prodromal AD from "normal aging" with BA35 measurements, whereas 52 per group (104 total) would be required with whole hippocampal volume.

Although this is the first study to comprehensively examine MTL subregional change across varying degrees of clinical progression, other work has supported the general principle of enhanced sensitivity to MCI with measurement of specific subregions (Iglesias et al., 2015; La Joie et al., 2013; Mueller et al., 2010). For example, La Joie et al. reported stronger group effects for CA1 than other subfields or volume of entire hippocampal formation in a manual tracing study (La Joie et al., 2013). Similarly, Mueller et al. found a CA1-2 transition zone to most strongly distinguish these groups (Mueller et al., 2010). In the current analysis, we also found CA1 to be the hippocampal subfield that tended to be associated with the most atrophy in MCI and AD. Notably, the above studies did not include PRC measurements, which would be expected to precede CA1 involvement. Indeed, in a different MCI cohort from here, we

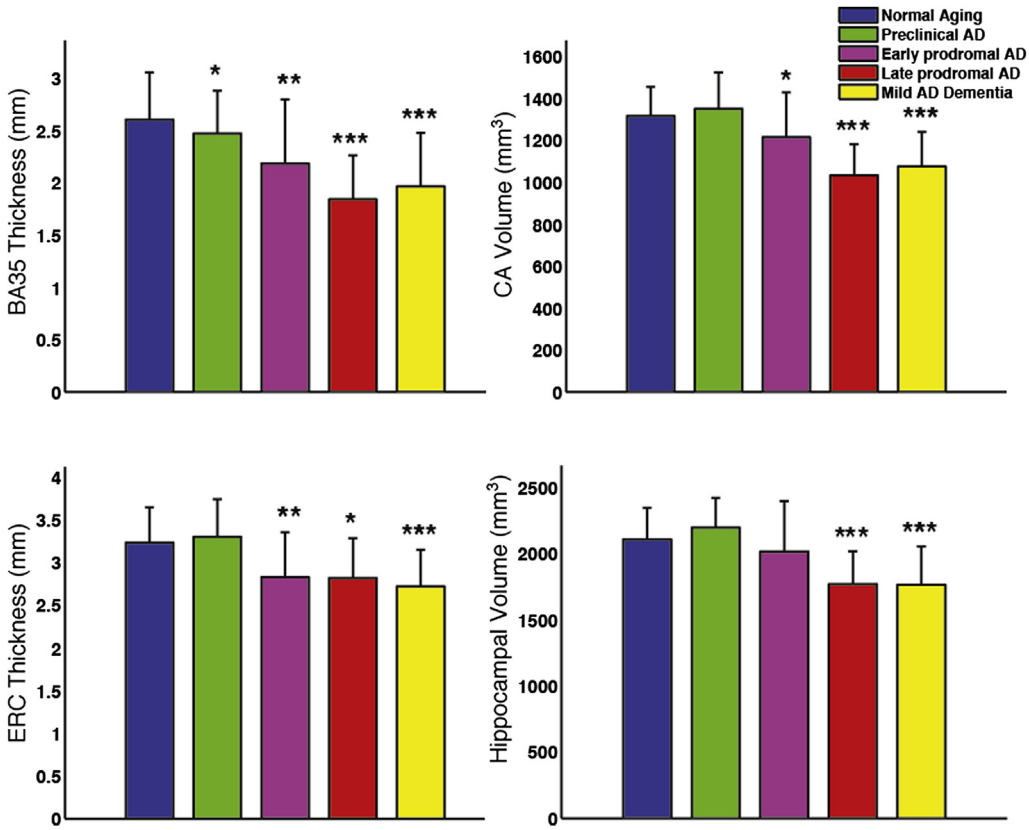

Fig. 2. Group comparison of select MTL subregional measurements adjusted for age, as well as ICV when volume measurements. Statistical significance is based on comparison with the "normal aging" (CN Am−) group. Error bars reflect 1 standard deviation. $^{*}p < 0.05$; $^{**}p < 0.01$; and $^{***}p < 0.001$. Abbreviations: AD, Alzheimer's disease; CN, cognitively normal adults; ERC, entorhinal cortex; MTL, medial temporal lobe.

reported strongest group discrimination with CN adults for both CA1 and BA35 (Pluta et al., 2012; Yushkevich et al., 2015). Although beyond the scope of the current analysis, it is also worth noting that the BA35 thinning found in the preclinical group also appeared to relate to performance on Auditory verbal learning test-delayed memory measures (5-minute delayed recall: $r = 0.52$, $p = 0.014$; recognition hits minus false alarms: $r = 0.51$, $p = 0.015$), supporting potential early cognitive consequences of this subtle atrophy.

Advantages of subregions relative to more global measures of hippocampal volume are less evident in later clinical stages, which is not unexpected given the relatively more diffuse involvement of MTL structures with disease progression (Braak and Braak, 1991). Nonetheless, examination of absolute differences in individual subregions suggests that there still may remain some degree of selectivity, but without translating into greater statistical significance. The increased precision of measurement of the entire hippocampus versus that of smaller subregions, particularly in later disease stages, likely mitigates against the value of the larger absolute differences detected in these regions.

4.2. Measurement of perirhinal cortex

There is relatively little work on in vivo measurement of PRC despite the medial portion of this structure, the "transentorhinal" region, being associated with the earliest NFT pathology (Braak and Braak, 1991). A number of reasons likely account for this limited attention, including anatomic variability, disagreement about boundaries, and even nosology, as there is inconsistency in whether the PRC is defined as including both BA35 and BA36 (Augustinack et al., 2013; Ding and Van Hoesen, 2010). Nonetheless, Augustinack et al. recently developed an automated surface-based labeling method of BA35 derived from annotation of very high-resolution postmortem scans (Augustinack et al., 2013). They applied this approach to ADNI data and also found very strong group discrimination between MCI patients and controls.

We used an in vivo atlas in which BA35 and BA36 were separately labeled and guided by description of the anatomic variability of these structures defined by sulcal patterns reported in the exhaustive histological work of Ding and Van Hoesen (Ding and Van Hoesen, 2010). Our prior work has suggested good reliability of automatic segmentations with manual ones (Yushkevich et al., 2015). Errors are largely driven by differences in the number of slices labeled anterior to the head of the hippocampus, as the atlas protocol artificially limits these segmentations to just 1 slice rostral to the hippocampal head. As such, the entire structure is not included, and the anterior extent is dependent on the size of the hippocampus, which is further influenced by disease status. This produces confounds for measurement of volume, which, for example, may be underestimated in an individual with a smaller hippocampal head. In light of this issue and some degree of uncertainty in the medial and lateral boundaries of these regions using sulcal pattern heuristics (see Fig. 3) that may not completely conform to cytoarchitecture, we measured mean cortical thickness, which likely limits the impact of these errors relative to volume measurements.

In prior work, we have also used thickness measures or "normalized" volumes based on the number of slices included in the segmentation, essentially a measure of area (Yushkevich et al., 2015). However, the latter does not account for uncertainty in the lateral boundaries or the fact that individuals with the same overall volume may differ along the anterior-posterior and medial-lateral dimensions. Notably, we previously found evidence of strong

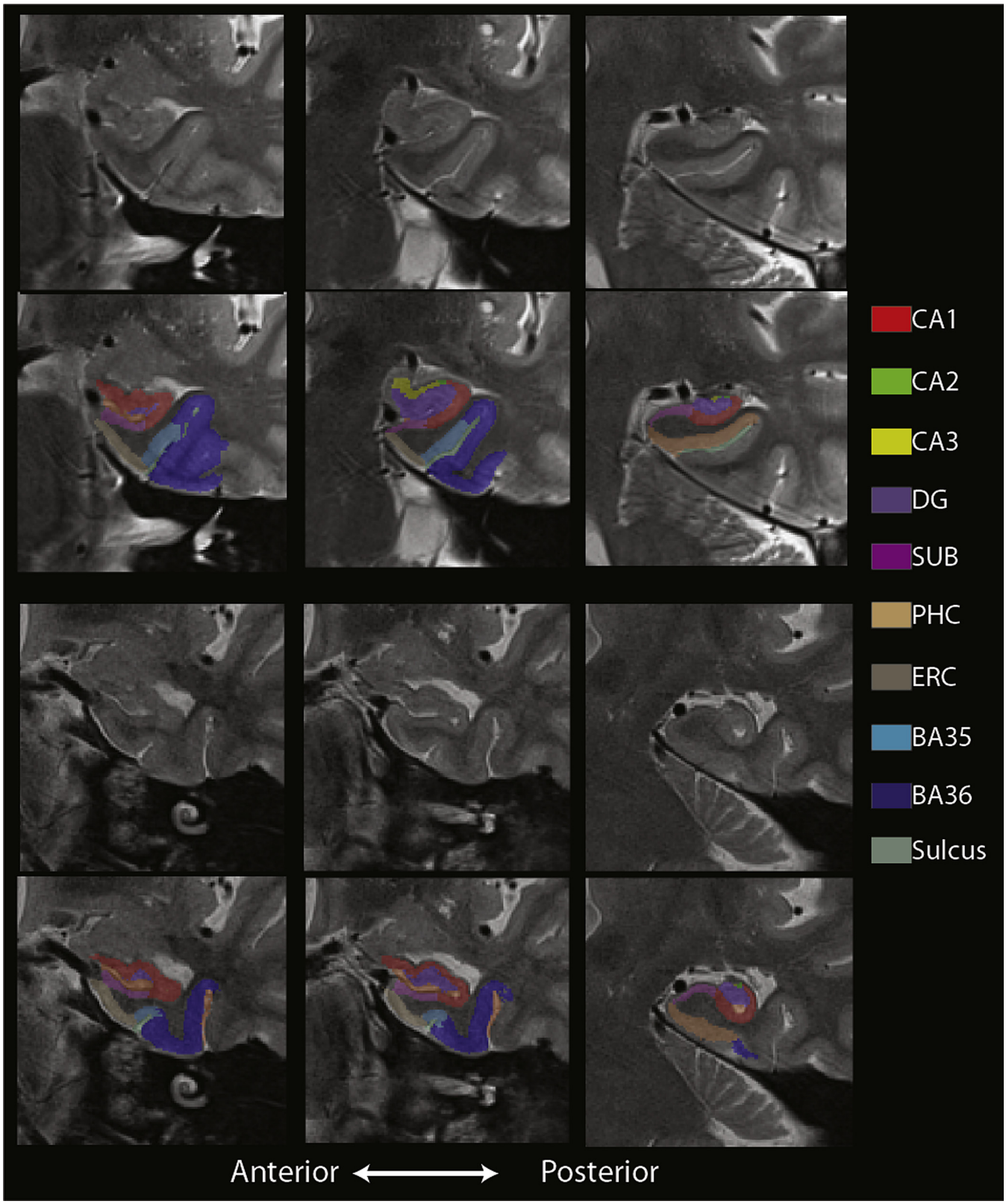

Fig. 3. Example of automatic segmentation in 2 individuals at several slices along the anterior-posterior axis. In particular, these cases exhibit the heuristic rules and variability of lateral boundaries in BA35 and BA36 based on collateral sulcus depth (top 2 rows: deep collateral sulcus; bottom 2 rows: shallow collateral sulcus). Abbreviations: CA1/2/3, cornu ammonis 1/2/3; DG, dentate gyrus; ERC, entorhinal cortex; PHC, parahippocampal cortex.

discrimination between MCI and controls in BA35 using thickness in a single-site study but used a T2-derived measure of thickness (see Yushkevich et al., 2015) rather than the T1-extracted thickness used here. The anisotropic nature and susceptibility to motion artifact of the T2 scans make this approach challenging, particularly in the ADNI data set, which is somewhat noisier for this sequence relative to that obtained in our prior study (Yushkevich et al., 2015). We did examine both methods in the current data set and found the T1 data more sensitive. We continue to work to improve these measurements both with regard to acquisition quality and analysis approach (Xie et al., 2014).

4.3. Pattern of atrophy recapitulates Braak staging

In addition to examining group statistics, we also calculated absolute percent change in volume and thickness across the spectrum of AD relative to "normal aging". This cross-sectional, pseudolongitudinal analysis is akin to the inference that Braak and Braak originally used to stage temporal ordering of NFT deposition in autopsy data (Braak and Braak, 1991). Our findings do appear to qualitatively match the expected pattern of regional involvement. Specifically, BA35 displayed the greatest percent difference (8.1%) in preclinical AD, whereas all other regions were less than 1% smaller or even slightly larger than the control group. The only exception was PHC, which was 5.2% thinner. In early prodromal AD, ERC and other extrahippocampal MTL regions displayed more significant cortical thinning, whereas hippocampal regions were more modestly affected. By late prodromal AD and mild AD dementia, hippocampal subfields were more prominently atrophied with CA being the most affected although the degree of absolute volume change was only marginally more than DG and subiculum. These findings broadly match the spread of NFT's initiated in transentorhinal cortex (BA35) and spreading to adjacent cortical MTL areas and then into the hippocampus proper focused in CA1 (Miller et al., 2015). However, one should be cautious in overinterpretation of these findings given their qualitative nature.

Interestingly, we found the BA36 portion of PRC displayed similar or better discrimination than BA35 in early prodromal AD. This result is somewhat unexpected given the greater NFT burden earlier in the latter. Despite being frequently grouped together as PRC, these 2 regions differ in cytoarchitecture with BA35 and 36 being best classified as periallocortex-proisocortex and isocortex, respectively, (Augustinack et al., 2013; Ding and Van Hoesen, 2010). It is worth noting that a more precise timing of BA36 involvement relative to BA35 is lacking in the literature, but clearly both regions are significantly affected in later symptomatic disease (Arnold et al., 1991; Braak and Braak, 1991; Van Hoesen et al., 2000). Importantly, BA35 displayed a larger degree of cortical thinning in absolute terms compared with BA36 across the spectrum from preclinical AD to mild AD dementia. The small size of BA35 and its location along the collateral sulcus likely contributes to increased variability, which may reduce its discriminatory power relative to BA36.

4.4. Limitations

Although this represents the largest cohort of MCI/AD patients using this MTL-specific T2-weighted MRI acquisition, the sample size remains modest, particularly when groups were divided based on amyloid status. This limits inference about the degree to which subregions are dissociable in their sensitivity to disease stage. Indeed, despite differences in apparent effect size, only BA35 significantly differed from whole hippocampal volume in discrimination of preclinical AD, whereas PHC displayed a trend toward such a difference in the early prodromal phase. It is also important to point out that a number of comparisons were made for each group comparison and not all significant effects survive Bonferroni correction. However, this analysis was not entirely exploratory and many of the reported findings are consistent with prediction based on Braak staging and the expected association with progressive atrophy in these regions. Further, the cross-sectional nature of the study makes arguments about sensitivity to disease progression tentative. For example, the fact that the groups are well-matched in age does not reflect the fact that in a longitudinal cohort both age and disease would contribute to change. Increased sample size and longitudinal analyses would also allow for more detailed examination of the predictive value of these measures and relationship to cognitive change.

The accuracy of thickness measurements in PRC, PHC, and ERC is limited by accuracy of gray matter segmentation in T1-MRI, as well as that of the respective ROI labels in T2-MRI. BA35, in particular, has lower segmentation reliability due to its small size and proximity to the collateral sulcus. Gray matter segmentation may also be affected by errors in resolving sulcal cerebrospinal fluid in this region. These factors introduce higher variance in the thickness measurements. However, these limitations also represent potential for improvements, and we continue to work on refined methodologic approaches to enhance reliability, such as additional strategies to measure thickness and combined use of T1- and T2-weighted sequences in the segmentation process (Xie et al., 2014; Yushkevich et al., 2015). Further, there are on-going collaborative efforts to use multimodal data, including histology, to develop a harmonized hippocampal subfield protocol, which has the potential to improve validity of these measures (hippocampalsubfields.com).

5. Conclusion

This article describes the use of an automated tool for segmentation of subregions within the MTL. Consistent with the selective topography of NFTs in early AD, subregional measurement appeared to offer enhanced sensitivity at preclinical and early prodromal disease stages compared with whole hippocampal volumes. As such, this may be a valuable biomarker for therapeutic trials in these populations.

Disclosure statement

Michael W. Weiner reports grants from Alzheimer's Association and General Electric; personal fees from Pfizer, Neurotrope BioScience, Alzheon, Inc, Eli Lilly, Synarc, Janssen, Avid Radiopharmaceuticals, Merck, Biogen Idec, BioClinica, Genentech, Nutricia, and Alzheimer's Drug Discovery Foundation. Susanne G. Mueller reports grants from the Alzheimer's Association. David A. Wolk reports grants from the Alzheimer's Association, Eli Lilly, Avid Radiopharmaceuticals, Merck, Biogen, and Functional Neuromodulation; personal fees from Piramal. The remaining authors have no conflicts of interest to disclose.

Acknowledgements

Funding support for this article included the following: National Institute on Aging; Contract grant number: P30AG010124 and R01 AG037376; National Institute of Biomedical Imaging and Bioengineering; Contract grant number: R01 EB014346. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug

Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References

Apostolova, L.G., Dinov, I.D., Dutton, R.A., Hayashi, K.M., Toga, A.W., Cummings, J.L., Thompson, P.M., 2006. 3D comparison of hippocampal atrophy in amnestic mild cognitive impairment and Alzheimer's disease. Brain 129 (Pt 11), 2867–2873.

Arnold, S.E., Hyman, B.T., Flory, J., Damasio, A.R., Van Hoesen, G.W., 1991. The topographical and neuroanatomical distribution of neurofibrillary tangles and neuritic plaques in the cerebral cortex of patients with Alzheimer's disease. Cereb. Cortex 1, 103–116.

Augustinack, J.C., Huber, K.E., Stevens, A.A., Roy, M., Frosch, M.P., van der Kouwe, A.J., Wald, L.L., Van Leemput, K., McKee, A.C., Fischl, B., 2013. Alzheimer's Disease Neuroimaging Initiative, Predicting the location of human perirhinal cortex, Brodmann's area 35, from MRI. Neuroimage 64, 32–42.

Braak, H., Braak, E., 1991. Neuropathological staging of Alzheimer-related changes. Acta Neuropathol. 82, 239–259.

Butters, N., Granholm, E., Salmon, D.P., Grant, I., Wolfe, J., 1987. Episodic and semantic memory: a comparison of amnesic and demented patients. J. Clin. Exp. Neuropsychol. 9, 479–497.

Das, S.R., Avants, B.B., Grossman, M., Gee, J.C., 2009. Registration based cortical thickness measurement. Neuroimage 45, 867–879.

Ding, S.L., Van Hoesen, G.W., 2010. Borders, extent, and topography of human perirhinal cortex as revealed using multiple modern neuroanatomical and pathological markers. Hum. Brain Mapp. 31, 1359–1379.

Fox, N.C., Warrington, E.K., Freeborough, P.A., Hartikainen, P., Kennedy, A.M., Stevens, J.M., Rossor, M.N., 1996. Presymptomatic hippocampal atrophy in Alzheimer's disease. A longitudinal MRI study. Brain 119 (Pt 6), 2001–2007.

Iglesias, J.E., Augustinack, J.C., Nguyen, K., Player, C.M., Player, A., Wright, M., Roy, N., Frosch, M.P., McKee, A.C., Wald, L.L., Fischl, B., Van Leemput, K., 2015. Alzheimer's Disease Neuroimaging Initiative, A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. Neuroimage 115, 117–137.

Jack Jr., C.R., Bernstein, M.A., Fox, N.C., Thompson, P.M., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging 27, 685–691.

Jack Jr., C.R., Petersen, R.C., O'Brien, P.C., Tangalos, E.G., 1992. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. Neurology 42, 183–188.

Jack Jr., C.R., Vemuri, P., Wiste, H.J., Weigand, S.D., Aisen, P.S., Trojanowski, J.Q., Shaw, L.M., Bernstein, M.A., Petersen, R.C., Weiner, M.W., Knopman, D.S., 2011. Alzheimer's Disease Neuroimaging Initiative, Evidence for ordering of Alzheimer disease biomarkers. Arch. Neurol. 68, 1526–1535.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156.

Kaplan, E., Goodglass, H., Weintraub, S., 1983. The Boston Naming Test. Lea and Feibiger, Philadelphia.

Kerchner, G.A., Hess, C.P., Hammond-Rosenbluth, K.E., Xu, D., Rabinovici, G.D., Kelley, D.A., Vigneron, D.B., Nelson, S.J., Miller, B.L., 2010. Hippocampal CA1 apical neuropil atrophy in mild Alzheimer disease visualized with 7-T MRI. Neurology 75, 1381–1387.

La Joie, R., Perrotin, A., de La Sayette, V., Egret, S., Doeuvre, L., Belliard, S., Eustache, F., Desgranges, B., Chételat, G., 2013. Hippocampal subfield volumetry in mild cognitive impairment, Alzheimer's disease and semantic dementia. Neuroimage Clin. 3, 155–162.

Landau, S.M., Mintun, M.A., Joshi, A.D., Koeppe, R.A., Petersen, R.C., Aisen, P.S., Weiner, M.W., Jagust, W.J., 2012. Alzheimer's Disease Neuroimaging Initiative, Amyloid deposition, hypometabolism, and longitudinal cognitive decline. Ann. Neurol. 72, 578–586.

Leow, A.D., Klunder, A.D., Jack Jr., C.R., Toga, A.W., Dale, A.M., Bernstein, M.A., Britson, P.J., Gunter, J.L., Ward, C.P., Whitwell, J.L., Borowski, B.J., Fleisher, A.S., Fox, N.C., Harvey, D., Kornak, J., Schuff, N., Studholme, C., Alexander, G.E., Weiner, M.W., Thompson, P.M., 2006. ADNI Preparatory Phase Study, Longitudinal stability of MRI for mapping brain change using tensor-based morphometry. NeuroImage 31, 627–640.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. Neurology 34, 285–297.

Miller, M.I., Ratnanather, J.T., Tward, D.J., Brown, T., Lee, D.S., Ketcha, M., Mori, K., Wang, M.C., Mori, S., Albert, M.S., Younes, L., 2015. BIOCARD Research Team, Network neurodegeneration in Alzheimer's disease via MRI based shape diffeomorphometry and high-field atlasing. Front. Bioeng. Biotechnol. 3, 54.

Morris, J.C., 1993. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 43, 2412–2414.

Mueller, S., Yushkevich, P., Wang, L., Van Leemput, K., Mezher, A., Das, S., Weiner, M., 2013. Collaboration for a systematic comparison of different techniques to measure subfield 1358 volumes: announcement and first results. Alzheimers Dement. 9, 51.

Mueller, S.G., Schuff, N., Raptentsetsang, S., Elman, J., Weiner, M.W., 2008. Selective effect of Apo e4 on CA3 and dentate in normal aging and Alzheimer's disease using high resolution MRI at 4 T. Neuroimage 42, 42–48.

Mueller, S.G., Schuff, N., Yaffe, K., Madison, C., Miller, B., Weiner, M.W., 2010. Hippocampal atrophy patterns in mild cognitive impairment and Alzheimer's disease. Hum. Brain Mapp. 31, 1339–1347.

Mueller, S.G., Stables, L., Du, A.T., Schuff, N., Truran, D., Cashdollar, N., Weiner, M.W., 2007. Measurement of hippocampal subfields and age-related changes with high resolution MRI at 4T. Neurobiol. Aging 28, 719–726.

Mueller, S.G., Weiner, M.W., 2009. Selective effect of age, Apo e4, and Alzheimer's disease on hippocampal subfields. Hippocampus 19, 558–564.

Petersen, R.C., 2004. Mild cognitive impairment as a diagnostic entity. J. Intern. Med. 256, 183–194.

Pluta, J., Yushkevich, P., Das, S., Wolk, D., 2012. In vivo analysis of hippocampal subfield atrophy in mild cognitive impairment via semi-automatic segmentation of T2-weighted MRI. J. Alzheimers Dis. 31, 85–99.

Reitan, R., 1958. Validity of the trail making test as an indicator of organic brain disease. Percept. Mot. Skills 8, 271–276.

Rey, A., 1964. L'examen clinique en psychologie Paris. Presses Universitaires de France, Paris.

Saykin, A.J., Wishart, H.A., Rabin, L.A., Santulli, R.B., Flashman, L.A., West, J.D., McHugh, T.L., Mamourian, A.C., 2006. Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. Neurology 67, 834–842.

Shing, Y.L., Rodrigue, K.M., Kennedy, K.M., Fandakova, Y., Bodammer, N., Werkle-Bergner, M., Lindenberger, U., Raz, N., 2011. Hippocampal subfield volumes: age, vascular risk, and correlation with associative memory. Front. Aging Neurosci. 3, 2.

Squire, L.R., 2004. Memory systems of the brain: a brief history and current perspective. Neurobiol. Learn. Mem. 82, 171–177.

Tustison, N.J., Cook, P.A., Klein, A., Song, G., Das, S.R., Duda, J.T., Kandel, B.M., van Strien, N., Stone, J.R., Gee, J.C., Avants, B.B., 2014. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. Neuroimage 99, 166–179.

Van Hoesen, G.W., Augustinack, J.C., Dierking, J., Redman, S.J., Thangavel, R., 2000. The parahippocampal gyrus in Alzheimer's disease. Clinical and preclinical neuroanatomical correlates. Ann. N. Y. Acad. Sci. 911, 254–274.

Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L.L., Augustinack, J., Dickerson, B.C., Golland, P., Fischl, B., 2009. Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI. Hippocampus 19, 549–557.

Wang, H., Yushkevich, P.A., 2013. Multi-atlas segmentation without registration: a supervoxel-based approach. Med. Image Comput. Comput. Assist. Interv. 16 (Pt 3), 535–542.

Wechsler, D., 1987. WMS-r Wechsler Memory Scale—Revised Manual. The Psychological Corporation, Harcourt Brace Jovanovich, Inc, New York.

Wisse, L.E., Gerritsen, L., Zwanenburg, J.J., Kuijf, H.J., Luijten, P.R., Biessels, G.J., Geerlings, M.I., 2012. Subfields of the hippocampal formation at 7 T MRI: in vivo volumetric assessment. Neuroimage 61, 1043–1049.

Xie, L., Pluta, J., Wang, H., Das, S.R., Mancuso, L., Kliot, D., Avants, B.B., Ding, S.L., Wolk, D.A., Yushkevich, P.A., 2014. Automatic clustering and thickness measurement of anatomical variants of the human perirhinal cortex. Med. Image Comput. Comput. Assist. Interv. 17 (Pt 3), 81–88.

Yushkevich, P.A., Pluta, J.B., Wang, H., Xie, L., Ding, S.L., Gertje, E.C., Mancuso, L., Kliot, D., Das, S.R., Wolk, D.A., 2015. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. Hum. Brain Mapp. 36, 258–287.

Yushkevich, P.A., Wang, H., Pluta, J., Das, S.R., Craige, C., Avants, B.B., Weiner, M.W., Mueller, S., 2010. Nearly automatic segmentation of hippocampal subfields in in vivo focal T2-weighted MRI. Neuroimage 53, 1208–1224.

Zeineh, M.M., Engel, S.A., Thompson, P.M., Bookheimer, S.Y., 2003. Dynamics of the hippocampus during encoding and retrieval of face-name pairs. Science 299, 577–580.