HHS Public Access

Author manuscript

Hum Brain Mapp. Author manuscript; available in PMC 2019 August 16.

Published in final edited form as:

Hum Brain Mapp. 2019 August 15; 40(12): 3431–3451. doi:10.1002/hbm.24607.

Automated segmentation of medial temporal lobe subregions on in vivo T1-weighted MRI in early stages of Alzheimer's disease

Long Xie[1,2], **Laura E. M. Wisse**[1,2,3], **John Pluta**[1,2], **Robin de Flores**[3,4], **Virgine Piskin**[1], **Jose V. Manjón**[5], **Hongzhi Wang**[6], **Sandhitsu R. Das**[1,3,4], **Song-Lin Ding**[7,8], **David A. Wolk**[3,4], **Paul A. Yushkevich**[1,2], **Alzheimer's Disease Neuroimaging Initiative**[†]

[1]Penn Image Computing and Science Laboratory (PICSL), Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania

[2]Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania

[3]Penn Memory Center, University of Pennsylvania, Philadelphia, Pennsylvania

[4]Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania

[5]Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universidad Politécnica de Valencia, Valencia, Spain

[6]IBM Almaden Research Center, San Jose, California

[7]Allen Institute for Brain Science, Seattle, Washington

[8]Institute of Neuroscience, School of Basic Medical Sciences, Guangzhou Medical University, Guangzhou, People's Republic of China

Abstract

Medial temporal lobe (MTL) substructures are the earliest regions affected by neurofibrillary tangle pathology—and thus are promising biomarkers for Alzheimer's disease (AD). However, automatic segmentation of the MTL using only T1-weighted (T1w) magnetic resonance imaging (MRI) is challenging due to the large anatomical variability of the MTL cortex and the confound of the dura mater, which is commonly segmented as gray matter by state-of-the-art algorithms because they have similar intensity in T1w MRI. To address these challenges, we developed a novel atlas set, consisting of 15 cognitively normal older adults and 14 patients with mild cognitive impairment with a label explicitly assigned to the dura, that can be used by the multiatlas automated pipeline (Automatic Segmentation of Hippocampal Subfields [ASHS-T1]) for the

Correspondence: Long Xie, Penn Image Computing and Science Laboratory (PICSL), 3700 Hamiton Walk, Richards building 6 floor, Philadelphia, PA 19104, long.xie@uphs.upenn.edu.

Long Xie and Laura E. M. Wisse contributed equally to this study.

[†]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.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

DATA ACCESSIBILITY

The data that support the findings of this study are openly available in NITRC at https://www.nitrc.org/frs/shownotes.php?release_id=3851 and ADNI database at http://adni.loni.usc.edu/.

segmentation of MTL subregions, including anterior/posterior hippocampus, entorhinal cortex (ERC), Brodmann areas (BA) 35 and 36, and parahippocampal cortex on T1w MRI. Cross-validation experiments indicated good segmentation accuracy of ASHS-T1 and that the dura can be reliably separated from the cortex (6.5% mislabeled as gray matter). Conversely, FreeSurfer segmented majority of the dura mater (62.4%) as gray matter and the degree of dura mislabeling decreased with increasing disease severity. To evaluate its clinical utility, we applied the pipeline to T1w images of 663 ADNI subjects and significant volume/thickness loss is observed in BA35, ERC, and posterior hippocampus in early prodromal AD and all subregions at later stages. As such, the publicly available new atlas and ASHS-T1 could have important utility in the early diagnosis and monitoring of AD and enhancing brain-behavior studies of these regions.

Keywords

Alzheimer's disease; anterior and posterior hippocampus; biomarker; dura mater; entorhinal cortex; mild cognitive impairment; perirhinal cortex; segmentation; T1-weighted magnetic resonance imaging; transentorhinal cortex

1 | INTRODUCTION

The medial temporal lobe (MTL) is the site of several neurodegenerative pathologies, most notably of neurofibrillary tangle (NFT) pathology, a hallmark of Alzheimer's disease (AD), which is thought to first affect the transentorhinal cortex, before it spreads to the entorhinal cortex (ERC) and cornu ammonis region 1 of the hippocampus (Braak & Braak, 1995, 1991; Ding, Van Hoesen, Cassell, & Poremba, 2009). As NFT pathology is closely related to neuron and synapse loss (Bobinski et al., 1997; Braak & Braak, 1991; Fukutani et al., 1995), certain MTL subregions may therefore show early and selective atrophy and serve as imaging biomarker in the early stages of AD. In fact, a recent in vivo magnetic resonance imaging (MRI) study showed selective atrophy in Brodmann area 35 (BA35), a region that approximates the transentorhinal region, in individuals with preclinical AD compared to controls (Wolk et al., 2017). These subregions are also of interest because they are thought to subserve different cognitive functions, such as recollection and familiarity (Wolk & Dickerson, 2011; Yonelinas et al., 2007), and are part of two dissociable MTL networks, where the anterior hippocampus, ERC and perirhinal cortex (PRC) are part of the anterior MTL network and the posterior hippocampus and parahippocampal cortex (PHC) are part of the posterior MTL network (Ranganath & Ritchey, 2012). These networks are also thought to be affected in the early stages of AD (Das et al., 2014).

Fine-grained measurement of subregions of the MTL has therefore received increasing attention in the recent years, with many studies utilizing high-resolution T2-weighted (T2w) MRI images, often with ~0.4×0.4 mm^2 in-plane resolution (Bender et al., 2018; Ekstrom et al., 2009; Mueller et al., 2007; Olsen et al., 2013; Preston et al., 2010; Yushkevich, Amaral, et al., 2015; Zeineh, Engel, Thompson, & Bookheimer, 2001). The advantage of these images is that they allow for improved visualization of MTL structures less visible on T1-weighted (T1w) MRI scans. For example, the stratum radiatum lacunosum moleculare, an important boundary between certain subfields of the hippocampus, and dura mater, which is

part of the meninges, can be visualized with T2w scans (Figure 1). The advantage of the clear visualization of the dura with these T2w MRI images is that it allows for accurate segmentation of important adjacent MTL subregions in contrast to T1w MRI images in which the dura has similar intensity as gray matter (Xie et al., 2016).

In our prior work (Yushkevich, Pluta, et al., 2015), we have developed a multiatlas segmentation software/package, called Automatic Segmentation of Hippocampal Subfields (ASHS, https://www.nitrc. org/projects/ashs), that can reliably segment MTL subregions in T2w MRI (the pipeline that works with T2w MRI is referred to as ASHS-T2 in this article). Even though there are advantages of these T2w MRI images over T1w MRI images, there are large data sets of T1w MRI scans available and analyzing these data sets would allow for more power to test hypotheses of interest. Additionally, T1w images often have higher resolution in the through-plane direction which helps in better resolving the folding and branching of sulci, important for the segmentation of these MTL cortical regions.

Available methods for the parcellation of MTL subregions on T1w MRI include several manual approaches (Kivisaari, Probst, & Taylor, 2013; Malykhin, Bouchard, Camicioli, & Coupland, 2008). An advantage of these manual approaches is that they often take into account the anatomical variability of the MTL cortex, which has multiple distinct anatomical subtypes, defined by the folding and branching patterns of the collateral sulcus (CS), that greatly affects the location of the borders between MTL cortices (Ding & Van Hoesen, 2010), for example, when the CS is deep, BA35 is located at the medial bank of the CS while when CS is shallow, BA35 also occupies the fundus and lateral bank. However, manual segmentation is infeasible for larger data sets like the AD neuroimaging initiative (ADNI), which includes thousands of MRI scans. Among the automated methods available for MTL subregion segmentation on T1w MRI, the specialized modules for MTL provided by FreeSurfer (Fischl, 2012) are of the most widely used in the literature in older populations (Delli Pizzi et al., 2016; Lehmann et al., 2010; Mah, Binns, & Steffens, 2015; Mishra et al., 2017; Pasquini et al., 2016). FreeSurfer includes a module for labeling hippocampal subfields and hippocampal lamina based on an ex vivo atlas (Iglesias et al., 2015). However, we have previously argued that standard resolution T1w MRI scans do not provide sufficient resolution for the visualization of the inner structure of the hippocampus and the parcellation of the hippocampal subfields (de Flores et al., 2015; Wisse, Biessels & Geerlings, 2014). FreeSurfer also provides separate specialized modules for ERC (Fischl et al., 2009) and PRC segmentation (Augustinack et al., 2013) based on spatial probability maps derived from ex vivo imaging. However, since these probability maps are defined in the space of a single template, this approach does not directly account for different subtypes of the MTL cortex.

Another important issue for T1w MRI scans, as mentioned above, relates to the visualization of the dura mater. In the MTL, a large proportion of the ERC and parts of the PRC are located directly adjacent to the dura and often appear merged with parts of the dura in T1w MRI (red arrows in Figure 1). To the best of our knowledge, none of the automatic analysis pipelines for MTL cortices using T1w MRI alone have addressed this confound, and the dura is often segmented as part of the gray matter by the state-of-the-art image processing algorithms, including FreeSurfer (the third column in Figure 1). This likely leads to an error

in the quantification of ERC and PRC, which potentially confounds the findings of research studies. In healthy individuals for whom there is little or no cerebrospinal fluid (CSF) separating the dura from the cortex, parts of the dura may be mistakenly labeled as cortex, whereas in patient groups with severe gray matter atrophy, the presence of such CSF would lead automatic methods to correctly exclude dura from the cortex label. This would potentially lead to a systematic bias in the estimation of group differences. To correct for this error, some studies performed manual correction either based on empirical rules or using T2w MRI of the same subject (e.g., the Human Connectome Project; Glasser et al., 2013). However, manual correction is labor intensive and T2w MRI scans are not always available. This is therefore not a feasible solution for large data sets consisting of only T1w MRI scans.

We hypothesize that the MTL cortex can be reliably automatically separated from the dura mater only using the T1w MRI even though there is only limited contrast between them. There are important features in T1w MRI that could be informative of the boundary of the dura and the cortex, for example, (a) the thin layer of CSF between the dura and the cortex can be visualized in some subjects (green arrow in Figure 1–b1) and (b) there are portions of the dura near the brain stem and inferior to the sulcus that are not merged with the cortex (white arrows in Figure 1). These features make it possible to infer the boundary between the cortex and the dura, even if it is not completely visible. Indeed, we recently developed a new atlas set that can be used by an established multiatlas segmentation framework (Yushkevich, Pluta, et al., 2015) together with a superresolution (SR) technique (Manjón et al., 2010) that is able to reliably segment ERC and PRC in T1w MRI that explicitly accounts for the confound of the dura mater and for the existence of multiple MTL cortex anatomical variants (Xie et al., 2016). To account for dura confounds, the atlas set for this pipeline was created using pairs of T1w and T2w scans in the same subjects, with the T2w previously used as an atlas set in the T2w-based MTL subregion segmentation approach by Yushkevich,Pluta, et al. (2015). The T2w images in the atlas set include labels for the ERC and subdivisions of the PRC, that is, BA35 and BA36, generated based on a segmentation protocol that takes anatomical variability of the CS into account and was developed in consultation with the neuroanatomist SLD, who has specific expertise in anatomical variability of the CS (Ding & Van Hoesen, 2010). The segmentations of this atlas set were transformed into the T1 space after coregistration with the T2w MRI of the same subject and manually edited to account for any residual misregistration errors. Additionally, the T1w atlas set segmentations were extended with a dura label informed by the aligned T2w MRI of the same subject, which helps in accurately locating the boundary between dura and the MTL cortex. Evaluation of this pipeline indicated that a large portion of the dura was assigned the correct label in our pipeline but not in other methods [FreeSurfer (Fischl, 2012) and ANTs (Avants, Epstein, Grossman, & Gee, 2008)], which included a large portion of the dura in the gray matter label. Cross-validation experiments showed promising segmentation accuracy of the automatic segmentation relative to manual segmentation for the MTL cortical regions [Dice similarity coefficient (DSC; Dice, 1945) is close to that using the T2w MRI (the ASHS-T2 pipeline, range from 0.671 to 0.755), in which the boundary between the cortex and the dura can be visualized]. Moreover, the clinical utility of the pipeline was evaluated by examining the statistical power in discriminating controls from amnestic mild

cognitive impairment (aMCI) patients, and indicated that BA35, in absolute terms, had the greatest area under the curve among the MTL cortex subregions, which is consistent with the Braak staging in the MTL (Braak & Braak, 1995).

This article extends this recent work, which was published in conference proceedings, with a richer set of MTL subregion measurements and additional experiments. We have extended our label set to include the PHC and the hippocampus, including a subdivision of the anterior and posterior hippocampus. Also, we provided thickness values in addition to volumes for the MTL cortices because thickness measures are less sensitive to border placement which may make them less sensitive to likely one aspect of errors in segmentation. We have improved the registration between the T1w and T2w MRI scans allowing for a closer alignment which required less editing of the transformed segmentations in the T1w-space. We performed crossvalidation experiments to assess the accuracy of the automatic segmentation relative to manual one, and compared our pipeline with FreeSurfer version 6.0 (Fischl, 2012) to evaluate how the different methods label dura in T1w MRI. And we further evaluated the performance of our pipeline in scans from ADNI phases GO and 2 by comparing MTL subregional volumes and thickness in amyloid negative controls with individuals with preclinical AD, prodromal AD and AD dementia. In addition, the atlas and software developed in this article are made publicly available (https://sites.google.com/view/ashs-dox/home). Finally, we have also provided an easy-to-use cloud-based service built into the ITK-SNAP image segmentation tool (Yushkevich et al., 2006) that allows users to execute our pipeline on a remote server with a few mouse clicks. The cloud based serviced is briefly summarized in Supplementary Material A. A Detailed tutorial of our cloud-based service is available at https://sites.google.com/view/ashs-dox/cloud-ashs/overview.

2 | METHODS

2.1 | Participants

2.1.1 | Penn Memory Center atlas set—The atlas set used in this study consists of 15 cognitively normal older controls (NC) and 14 aMCI patients. These participants were recruited from the Penn Memory Center/Alzheimer's Disease Center (PMC/ADC) at the University of Pennsylvania. Diagnosis of aMCI was made following established criteria (Petersen, 2004; Petersen et al., 2009; Winblad et al., 2004). This study was approved by the Institutional Review Board of the University of Pennsylvania and informed consent was provided by all subjects. This is the same atlas set that was used by Yushkevich, Pluta, et al. (2015) and Xie et al. (2017) to develop the atlas set using both T1w MRI and high-resolution T2w MRI. To avoid confusion, the atlas set developed in this study will be referred to as the **PMC-T1 atlas** and the one used in Yushkevich, Amaral, et al. (2015), Yushkevich, Pluta, et al. (2015), and Xie et al. (2017) will be referred to as the **PMC-T2 atlas**. Demographic and the mini-mental state examination (MMSE) data for the aMCI and NC groups are shown in Table 1.

2.1.2 | Data set from the ADNI—Part of the data used in the preparation of this article was obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003

as a public–private partnership, led by Principal Investigator Michael W. Weiner. 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. For up-to-date information, see www.adni-info.org.

Cognitively normal controls and amyloid-beta (Aβ) positive patients that have T1w MRI scans available from the ADNI GO and ADNI 2 were included in this study. The amyloid status of each participant is determined by thresholding a summary measure of Florbetapir standardized uptake value ratio (SUVR) derived from Florbetapir PET using a threshold of SUVR > = 1.11 (Landau et al., 2012). The summary Florbetapir SUVR measure came from publicly available processed data on the ADNI website, which was calculated by taking the mean SUVR of a set of regions typically associated with increased uptake in AD and using cerebellar gray matter as reference region [details described in Landau et al. (2012)]. In total, 667 participants were included and grouped into Aβ negative (Aβ−) controls, preclinical AD (Aβ positive controls), early prodromal AD (Aβ positive early MCI), late prodromal AD (Aβ positive late MCI) and dementia patients (Aβ positive AD). Four subjects' T1w MRI scans suffered from severe motion and thus were excluded from in this study. Table 2 summarizes the characteristics of the remaining 663 subjects.

2.2 | Neuroimaging data acquisition

2.2.1 | Imaging protocol of the PMC atlas set—The MRI scans of the atlas set were acquired on a 3T Siemens Trio MRI scanner (Erlangen, Germany) at the University of Pennsylvania using an eight-channel array coil. The imaging protocols include: (a) a whole brain T1w (magnetization prepared rapid acquisition gradient echo) MRI scan and (b) a T2w (turbo spin echo,) MRI scan with partial brain coverage and oblique coronal slice positioned orthogonally to the main axis of the hippocampus (De Vita et al., 2003; Thomas et al., 2004). The parameters of the T2w MRI are: TR/TE = 5310/68 ms, 18.3 ms echo spacing, 15 echo train length, 150°flip angle, 0% phase oversampling, 0.4 × 0.4 mm^2 in-plane resolution, 2.0 mm slice thickness with 0.6 mm gap, 30 interleaved slices, and 7:12 min acquisition time. For the T1w MRI, they are: TR/TE/TI = 1600/3.87/950 ms, 15° flip angle, 1.0 × 1.0 × 1.0 mm^3 isotropic resolution, and 5:13 min acquisition time.

2.2.2 | ADNI imaging protocol—The MRI imaging protocols of the ADNI study that were used to acquire the T1w MRI scans were previously described in Jack et al. (2008) and Leow et al. (2006). For Florbetapir PET, images were acquired in a 20 min PET brain scan session (four frames of 5 min duration) after a 50-min uptake phase following injection of 10 mCi of tracer.

2.3 | Manual segmentation of the MTL subregions in T1w MRI

The procedure of manual segmentation can be divided into two steps: manual segmentation of the MTL cortex and the hippocampus. Manual segmentation of the MTL cortex was initialized with the manual segmentations of the PMC-T2 atlas set (in the space of the T2w MRI) propagated to the space of the T1w MRI. Information from both T1w and T2w MRI scans of the same subject was taken into account during manual segmentation, which is

crucial for separating dura from the cortex. For the hippocampus, an automatic segmentation was first generated and followed by manual editing. All edits and segmentations were performed using ITK-SNAP (Yushkevich et al., 2006).

2.3.1 | Segmentation of MTL cortex and dura—Manual segmentations of the MTL cortex from the PMC-T2 atlas set from Yushkevich, Pluta, et al. (2015) and Xie et al. (2017) were propagated to the space of the aligned T1w MRI, followed by manual edits and addition of the dura label (more details regarding the segmentation protocol for the MTL cortex can be found in these two citations). Figure 2 shows three examples that illustrate the workflow. Details are described below.

Alignment between T2w MRI and the T1w MRI of the same subjects were performed following the steps below:

1. The T1w MRI was rigidly aligned to the T2w MRI using the ANTs (http://stnava.github.io/ANTs/) with mutual information as the similarity metric.
2. The T1w MRI was upsampled to $0.5 \times 0.5 \times 1.0\ mm^3$ by applying a patch-based SR technique (Manjón et al., 2010) for the purpose of bringing the resolution of the T1w MRI closer to that of the T2w MRI. Also, the SR upsampling increases the contrast between the dura and gray matter in T1w MRI so that the boundary between them can be better visualized.
3. The T2w MRI and the corresponding manual segmentation were resampled to $0.4 \times 0.4 \times 1.3\ mm^3$ using linear and nearest neighbor interpolation, respectively. The purpose of this step is to make the voxel size of the T2w MRI and SR T1w MRI similar.
4. The interpolated T2w MRI was cropped based on its manual segmentation with a margin of 10 voxels in all directions. This is done separately for left and right hemispheres.
5. From experiments, we found that global rigid intermodality registration [Step (1)] is not sufficient to accurately align the bilateral MTL regions which is probably due to small local spatial distortion of the two modalities. In order to generate better alignment of the MTL between the two modalities, for each hemisphere, affine registration was performed between the SR T1w MRI and the cropped upsampled T2w MRI, initialized with the rigid transformation between the whole brain T1w MRI and T2w MRI obtained in Step (1). This additional local affine registration, which was not included in the prior work (Xie et al., 2016), is essential for accurate alignment of the MTL between the two modalities.
6. The SR T1w MRI was transformed and resampled to the cropped upsampled T2w MRI space (referred to as registered SR T1w MRI), in which manual segmentations of the MTL cortex and the hippocampus were performed.

After registration of the T2w and T1w MRI described above, the MTL region of both modalities are well aligned as shown in the first two images of each example in Figure 2.

Labels of the MTL cortex, including cortical labels (ERC, BA35, BA36, PHC) and sulcus labels (CS and occipitotemporal sulcus [OTS]), were copied over to the registered SR T1w MRI (the third and the fourth images of each example in Figure 2). Because of slight differences in appearance between T1w and T2w MRI and small errors in registration due to highly anisotropic voxel size of T2w MRI, intermodality registration and the upsampling of both modalities, the labels were checked and manually edited to correctly match the border with the white matter, CSF and dura. For these edits, both the T1w and T2w MRI of the same subject were opened side by side in ITK-SNAP so that boundaries can be determined using information from both modalities (the fifth image of each example in Figure 2). Note that only the outer borders with surrounding regions were adjusted, not the borders between the different MTL cortices. Only the last slice of the ERC was adjusted, as a transition slice, extending half the length of one slice anterior (note that these two slices translate to one slice on the T2w MRI). This is similar to the procedure in Berron et al. (2017). Because of small registration errors and perhaps slight differences in the visualization of the hippocampus on T1w and T2w MRI, the anterior and posterior borders of the MTL cortices did not follow the protocol as described in Yushkevich, Pluta, et al. (2015) for some subjects. To reach consistency between subjects but to minimize changes to the original segmentations, an optimal anterior and posterior border was decided upon based on the full atlas set. The ERC, BA35, and BA36 extend one 1.3 mm slice anterior to the first slice of the hippocampus (was one 2.6 mm slice in the original protocol), ERC extends two 1.3 mm slices posterior to the most posterior slice of the uncus (same as in original protocol) and BA35/BA36 extends four 1.3 mm slices (same as in original protocol). The most anterior slice of the PHC is one slice posterior to the end of BA35 and BA36 (same as in original protocol) and the most posterior slice is fourth most posterior 1.3 mm slice of the hippocampus (was second most posterior 2.6 mm slice in original protocol). All subjects were visually checked and the segmentations were adjusted to match these boundary rules. Any given label needed to be extended at most two slices, where borders were matched to adjacent slices. In none of the cases the anatomy changes dramatically from one slice to the next, making these adjustments feasible.

Importantly, along the full length of the MTL cortex, a label for the dura mater was assigned to the voxels inferior to the corrected MTL cortex labels that have gray appearance in the registered SR T1w MRI and dark appearance in the resampled T2w MRI. Of note, the segmentation of the dura was informed by the registered T2w MRI, from which the boundary between dura and the cortex can more easily be identified. This is especially crucial for the situation when dura is completely attached to the cortex and is difficult to visualize in T1w MRI (Example 1 in Figure 2). In some cases, a thin layer of CSF between the dura and gray matter is visible in SR T1w MRI (green arrow in Example 2 in Figure 2), that is, a layer of voxels that have much darker intensity between the dura and gray matter in SR T1w MRI, which helps guide the dura segmentation. The CSF voxels were assigned a miscellaneous label. Moreover, in some cases, this layer of CSF is not visible; however, the dura is not completely attached to the cortex either (Example 3 in Figure 2). The portion of the dura near the brain stem and inferior to the CS that is not adjacent to the cortex (white arrows in Figure 2) also provides clues for automatic and manual segmentation of the dura. The anterior and posterior extents of the dura are limited to the slices with MTL cortex labels (ERC, BA35, BA36, and PHC).

2.3.2 | Segmentation of the hippocampus—

The European Alzheimer's Disease Consortium and ADNI harmonized protocol (HarP) (Boccardi et al., 2015; Frisoni et al., 2015) is a well-validated HarP for hippocampus segmentation in T1w MRI. To be consistent with the HarP protocol, we chose a subset (11 controls, 13 MCI patients, and 8 AD patients) of the publicly available HarP training set from ADNI described in Boccardi, Bocchetta, Morency, et al. (2015) as the training set to automatically segment the registered SR T1w MRI (obtained in Section 2.3.1) of the 29 cases in the PMC-T1 atlas set using the ASHS package/software (see Section 2.3.1 for a brief description). These automatic segmentations were used to initialize the manual segmentation of the hippocampus. The characteristics of the HarP training set and the details of the automatic segmentation pipeline are described in Supplementary Material B.

All segmentations were visually checked and edited where necessary in three planes, following the HarP protocol. Two adjustments were made to the HarP protocol. First, the medial border of the hippocampus was extended to be continuous with the MTL cortex generated in Section 2.3.1. The medial border therefore followed the protocol from Wisse et al. (2012). In the most posterior slices, the hippocampal medial border was located at the most medial point of the cortex not including the calcarine sulcus. The hippocampus did not always reach the parahippocampal gyrus in these most posterior slices. Second, this medial border was executed until the slice where the hippocampus was embedded in the splenium, to improve the transition to the decreasing size of the hippocampus on consecutive slices. Moreover, the hippocampus was split in an anterior and posterior region, where the border was defined by the most posterior slice of the uncus, which was included in the anterior hippocampus (Malykhin et al., 2007).

2.4 | Automatic segmentation using ASHS

2.4.1 | Construction of ASHS-T1 atlas using the ASHS training pipeline—

The original T1w MRI, whole-brain SR T1w MRI together with the bilateral manual segmentations in the space of the SR T1w MRI are fed into the ASHS training pipeline to generate an atlas (**ASHS-T1** atlas). The ASHS-T1 atlas (https://www.nitrc.org/frs/shownotes.php?release_id=3851) and the ASHS pipeline (https://www.nitrc.org/projects/ashs) are publicly available. The ASHS training pipeline is described in detail in Yushkevich, Pluta, et al. (2015) and summarized briefly in the following steps:

1. An unbiased whole brain population template is built using the T1w MRI of all the subjects.
2. The region of interest (ROI) of each hemisphere was identified by averaging the corresponding manual segmentations that are warped to the space of the template.
3. Each SR T1w MRI and the corresponding segmentation were warped to the space of the template and cropped around the ROI.
4. Pairwise registrations between all the subjects were performed between the warped and cropped scans.

5. Label fusion was performed for each atlas in its native space using the rest of the atlases as candidates.

6. AdaBoost classifiers were trained to learn the systematic error between the automatic segmentation and the manual segmentations.

2.4.2 | Application of ASHS-T1 atlas to new images—Once the ASHS-T1 atlas is trained, we can use the ASHS segmentation pipeline to automatically segment the T1w MRI scan of a new subject. Different from the pipeline described in Yushkevich, Pluta, et al. (2015), the proposed pipeline only takes the T1w MRI scan as input and does not require the T2w MRI scan. In brief, it involves the following steps:

1. The T1w MRI of the target subject is first upsampled to $0.5 \times 0.5 \times 1$ mm^3 using the SR technique (Manjón et al., 2010).

2. The ROI around the left and right MTL are identified in the target SR T1w image by registering to a whole-brain template generated in the training pipeline.

3. For each target ROI, the corresponding ROIs in the atlas set are registered to it using ANTs with normalized cross-correlation metric (Avants et al., 2008).

4. Atlas labels are then warped to the target ROI and combined using the joint label fusion algorithm (Wang & Yushkevich, 2013).

5. The process is repeated in a bootstrapping fashion, where the initial segmentation of the target structures is used to initialize affine alignment between the atlas and target ROIs. This bootstrapping results in fewer failed atlas-to-target registrations and better overall segmentation accuracy. The automatic segmentation generated from this step is referred to as the "**Heur**" output (The name "Heur" stands for heuristic rules that can be specified in ASHS. In our prior work on T2-weighted MRI (Yushkevich, Pluta, et al., 2015), we apply some heuristics to cut off anterior/posterior parts of cortical labels. Although no heuristic rules were used in this study, we keep the naming convention the same to be consistent with the outputs of the ASHS software.).

6. Two Adaboost classifiers, which were trained on shape features (the output referred to as the "**NoGray**") or shape and gray-scale intensity features ("**UseGray**") to correct for systematic errors generated in the multiatlas label fusion step, are applied to further improve the automatic segmentation. Since the classifiers were trained on the images of the atlas set, they may not generalize well to images acquired with different MRI imaging protocols. Therefore, using the "UseGray" output is only recommended if the target T1w MRI scan is acquired with a similar protocol as the atlas set.

Final bilateral automatic segmentations are generated in the target SR T1w MRI space. For the atlas set of 29 subjects, the automatic segmentation in the space of the SR T1w MRI was generated in a leave-one-out manner using the remaining 28 subjects as atlases. The segmentation accuracy of the "UseGray" output is reported in Table 3 and those of all the three outputs ("Heur", "NoGray," and "UseGray") are also computed and reported in Table S1.

When segmenting the baseline scans of the ADNI cohort, the whole 29-subject ASHS-T1 atlas set was used. Because of the difference in imaging protocol between the ADNI and the PMC-T1 atlas set, it is not appropriate to use the "UseGray" output. When quality control (Section 2.4.3) was performed, we observed that the dura and MTL cortex segmentation is of better quality when using the "Heur" output compared to the "NoGray" and thus the former is used. The segmentation accuracy in terms of DSC between automatic and manual segmentations of "Heur" is comparable but slightly lower (1.5% maximum DSC) than "UseGray" shown in Table S1. Volumetric and thickness (see Section 2.5.3) measurements of bilateral anterior/posterior hippocampus, ERC, BA35, BA36, and PHC were extracted for each subject.

2.4.3 | Quality control—The quality of all the automatic segmentations generated by ASHS-T1 was visually checked. The pipeline successfully labeled the baseline T1w MRI scans of all the 663 ADNI subjects while small errors occurred in a small subset of the subjects. Specifically, we observed undersegmentation in the lateral border of the hippocampus in 28 out of 663 subjects (seven Aβ− control, two preclinical AD, eight early prodromal AD, seven late prodromal AD, and four dementia). In 16 out of 663 subjects (one Aβ− control, one early prodromal AD, six late prodromal AD [one overlap], and eight dementia [one overlap]), oversegmentation of MTL cortices were identified. This is unavoidable partially due to the lack of contrast between cortex and dura. Examples of the common segmentation errors are shown Figure S1.

2.5 | Additional image processing

2.5.1 | Intracranial volume—Intracranial volume (ICV) was segmented from the T1w MRI of each ADNI subject using an in-house ICV segmentation software using ASHS with a training set of 27 T1w MRI scans (15 controls and 12 aMCI) and the corresponding manual ICV segmentations. The manual labels in this atlas set were generated with the guidance of the coregistered computer tomographic (CT) scans of the same subjects. Since the boundary between the skull and the soft tissue is clear in CT scans, we were able to obtain a relatively accurate manual segmentation of the ICV. Supplementary Material C describes the detail of ICV automatic segmentation pipeline.

2.5.2 | Cross-validation experiment in the atlas set in the space of the T2w MRI (ASHS-T2)—To compare the segmentation accuracy of the MTL cortices of the proposed pipeline that only utilizes T1w MRI to that using both T1w and T2w MRI (Yushkevich, Amaral, et al., 2015; Yushkevich, Pluta, et al., 2015), leave-one-out cross validation was also performed using the PMC-T2 atlas (comparisons were performed between the automatic and manual segmentations in the space of the T2w MRI). The same experiment has been done in Yushkevich, Pluta, et al. (2015). However, since we have updated the ASHS software (ASHS version 2.0.0 rather than 1.0.0, https://www.nitrc.org/frs/?group_id=370) and the atlas manual segmentation [the PHC and OTS labels were added as described in Xie et al. (2017)], the results are slightly different from that in Yushkevich, Pluta, et al. (2015). Note that we did not perform this analysis for the hippocampus, as the segmentation protocol for the T1w and T2w hippocampus were different.

2.5.3 | Thickness measures of the MTL cortices extracted from the ASHS-T1 automatic segmentation—

For MTL cortices, thickness measures may be more appropriate compared to volume because they are less sensitive to uncertainty in boundary estimation between cortical regions. A multitemplate thickness analysis pipeline (Xie et al., 2017, 2014) was applied to the MTL cortex labels (ERC, BA35, BA36, and PHC) to extract thickness. Since large anatomical variability, that is, different branching and folding patterns of the cortex, exists at the MTL cortex, traditional single-template-based approaches may not generate accurate thickness measures. The thickness pipeline takes anatomical variability into account by fitting corresponding variant-specific template to the target segmentation, which has been shown to generate more accurate thickness measurement (Xie et al., 2017, 2014).

2.5.4 | Volume and thickness measures of hippocampus, ERC, and PRC using FreeSurfer—

In order to compare the volume and thickness measurements extracted from the proposed pipeline to that from an established paradigm for T1w MRI, FreeSurfer version 6.0 (Fischl, 2012) was applied to the T1w MRI scans of both the 29 subjects in the atlas set and the ADNI data set. Volume measurements of the hippocampus were extracted from the "aseg.stats" file and volume and thickness measures of the ERC and PRC were extracted from "lh.BA_exvivo.thresh. stats" and "rh.BA_exvivo.thresh.stats" files. The location of the ERC and PRC is estimated using a probabilistic framework with templates constructed from ex vivo atlases described in Fischl et al. (2009) and Augustinack et al. (2013), respectively.

2.6 | Statistical analysis

All statistical analyses in this article are two-tailed with significance levels of $p = .05$ unless stated otherwise. Bilateral measurements of each subregion were averaged.

2.6.1 | Analysis of demographic and MMSE data—

To test the differences of demographic and MMSE between diagnosis groups, that is, aMCI-NC of the PMC atlas set and each patient-control pair of the ADNI data set, independent two-sample t test (age), Wilcoxon rank sum test (education, MMSE), and contingency χ^2 test (gender) were performed.

2.6.2 | Evaluate the accuracy of the automatic segmentation—

To evaluate the automated segmentations generated by ASHS-T1 and ASHS-T2, average DSC (Dice, 1945) between the leave-one-out automatic segmentations and the corresponding manual segmentations of each image in the PMC atlas sets were computed. In addition, we also computed the intraclass correlation (ICC) between volume measurements of the MTL subregions extracted from the automatic segmentations in the PMC-T1 atlas set and those obtained using the edited manual segmentations in T1w MRI space. To compare the ICC for the ASHS-T1 pipeline with that of the ASHS-T2 pipeline, similar analysis was performed for the MTL cortex labels (ERC, BA35, BA36, and PHC) for the PMC-T2 atlas set as well. ICC is computed using the "icc" function with the R package psy 1.1 implementing the method described in Shrout and Fleiss (1979).

2.6.3 | Group analysis between patients and Aβ−controls in ADNI—To evaluate the clinical utility, the four patient groups were compared to Aβ− controls separately. For each volume measure, a general linear model with group membership as the factor of interest, age and ICV as covariates, was fitted to obtain the t statistics for the control-patient contrast. Bonferroni corrected significance level ($p < .05/10$) is used to determine significant effects. For thickness measures, similar analysis was performed but only age was used as covariate and the Bonferroni correction significance level was set to $p < .05/6$.

3 | EVALUATION EXPERIMENTS AND RESULTS

We first evaluated the accuracy of the automatic segmentation of ASHS-T1 with the manual ones in the space of the T1w MRI and compare the performance of ASHS-T1 with that of ASHS-T2 (Section 3.1). Then, we investigated the extent to which an established analysis method for T1w MRI, that is, FreeSurfer, mislabels the dura mater and the cortex (Section 3.2). Finally, to demonstrate clinical utility of the proposed pipeline, we compared the volume and thickness measures extracted using the proposed pipeline between patients and controls using a large data set from the ADNI and compared this with FreeSurfer (Section 3.3).

3.1 | Evaluate accuracy of the automatic segmentation with manual segmentation

Primary validation of segmentation accuracy was performed on the set of 29 subjects from the PMC atlas for whom T1w MRI, T2w MRI, and both automatic and manual segmentations of the SR T1w MRI and T2w MRI are available.

The DSC results and volume measurements are summarized in Table 3. High DSCs of anterior (0.92), posterior (0.90), and whole (0.93) hippocampus segmentation were observed. The good accuracy in segmenting dura (0.75) and the MTL cortex labels (ERC: 0.76, BA35: 0.71, BA36: 0.79, and PHC: 0.80) indicates that the proposed pipeline can reliably segment the dura and the subregions of the MTL cortex. The slightly lower DSC of BA35 is not surprising given that it is a small structure and high anatomical variability exists in this region. No significant differences in segmentation accuracy, tested by two-sample *t* test, were found between aMCI and NC for all labels. In addition, the DSCs of the MTL cortex labels of ASHS-T1 are comparable to that of the ASHS-T2 (Yushkevich, Pluta, et al., 2015). The DSC of the proposed pipeline in segmenting ERC (DSC = 0.76) is slightly lower than that in T2w MRI (DSC = 0.79), which could be due to the limited ability to resolve gray matter boundaries because of the lower resolution and the confound of dura in T1w MRI. Importantly, we observe that the volume of the dura mater is larger than that of the ERC and BA35, indicating that segmenting dura mater as cortex could significantly confound volume measures of subregions in the MTL cortex.

From the ICC results, as reported in Table 3, ASHS-T1 demonstrates high accuracy in segmenting anterior/posterior hippocampus (0.95 and 0.89), BA35 (0.77), and BA36 (0.76). The ICC for ERC (0.69) and PHC (0.64) were slightly lower. The ICC values did not show notable differences between the ASHS-T1 and T2 pipeline. According to the Bland–Altman plots, shown in Figure 3, there exists a small bias in ERC segmentation, that is, the pipeline tends to undersegment larger ERC volumes and oversegment smaller ones. This would likely

lead to a slight underestimation of group differences for ERC. No bias is observed for the other subregions.

3.2 | Dura mislabeling as cortex

In this section, we performed experiments to test the two hypotheses that were introduced in the Introduction, that is, (a) the MTL cortex is commonly oversegmented by FreeSurfer because of the mislabeling of the adjacent dura mater and (b) the degree of dura mislabeling as cortex by FreeSurfer is different between patients and controls.

To test the first hypothesis, among subjects in the PMC-T1 atlas set, we first resampled the FreeSurfer whole brain segmentations to the space of the SR T1w MRI and then computed the average percentage of voxels labeled as dura in the manual segmentations that were mislabeled as gray matter or other by the proposed pipeline and FreeSurfer. The results, shown in Table 4, support the notion that a large proportion of dura (62.4%) is segmented as gray matter by FreeSurfer. We note that FreeSurfer does not have a specific label for the dura and therefore has to label the dura voxels as something else; including them in the gray matter introduces error to cortical thickness computations. On the other hand, the majority (71.9%) of dura voxels are correctly labeled by the proposed pipeline, only 6.5% of them are labeled as gray matter and the amount of dura mislabeling as cortex is not significantly different between aMCI and NC (6.8 ± 3.1% vs. 6.2 ± 4.2%, $p > .1$, revealed by two-sample t test).

The second hypothesis can be tested using the ADNI data set with controls and patients at different stages of AD. Since manual segmentation of the MTL cortex and dura is not available in the ADNI data set, the degree of dura mislabeling as gray matter by FreeSurfer is computed using the automatic ASHS-T1 segmentation, that is, the average percentage of voxels labeled as dura by the ASHS-T1 that are labeled as gray matter. We believe this is a suitable measure because of the following evidence:

1. In the PMC-T1 atlas set, we computed the degree of dura mislabeling as gray matter by FreeSurfer relative to the dura label in the automatic segmentations generated by ASHS-T1 and relative to the dura label from the manual segmentations. These measurements were highly correlated (Pearson correlation $r = .946$, $p = 9.3\ e{-}15$), shown in Figure S2.
2. In the PMC-T1 atlas set, no significant differences were observed between aMCI and controls in segmentation accuracy of dura (DSC reported in Table 3, 0.74 vs. 0.76) or for mislabeling of dura as cortex (6.8 vs. 6.2%) using the automatic dura segmentations generated by ASHS-T1. Therefore, it is unlikely that it will introduce bias between patients and controls.
3. All the segmentations of the ADNI subjects generated by the T1 pipeline used in this analysis were visually checked and only segmentations that have high-quality MTL cortex segmentation were used in this analysis and thus the bias induced by segmentation errors is limited.

Figure 4 summarizes the percentage of dura voxels segmented as gray matter by FreeSurfer in Aβ− controls and the four patient groups. The amount of dura mislabeling as cortex

decreases with increasing disease severity, probably due the more distinct separation between the MTL cortex and the dura (Figure 4). The proportion of mislabeling is significantly different between Aβ− controls and patients at early prodromal AD, late prodromal AD, and dementia stages revealed by two-sample *t* tests. Since manual segmentation of the ADNI data set is not available, it is not feasible to evaluate the amount of dura mislabeling as cortex of the proposed method. However, since we did not see large difference of dura mislabeling as cortex between aMCI and controls in the PMC-T1 atlas set (0.6%), it seems unlikely that the observed large differences of FreeSurfer dura mislabeling between groups (3.5, 6.5, and 8.6% between patients at early, late prodromal AD, dementia, and controls, respectively) are mainly due to imperfect automatic segmentation of ASHS-T1.

3.3 | MTL atrophy in early stages of AD in ADNI

We compared the volume and thickness measures extracted using the proposed pipeline between patients at different stages of AD and Aβ−controls in ADNI and performed a comparison with FreeSurfer. To make this comparison fair, we report here the ASHS-T1 results on the full ADNI data set, without excluding subjects based on the quality control procedure described in Section 2.4.3. However, excluding the subjects with poor ASHS-T1 segmentation quality did not significantly alter the comparison with FreeSurfer, as shown in Table S2.

As shown in Tables 5 and 6, we observed Bonferroni-corrected significant group effects at the early prodromal AD stage in posterior hippocampus volume ($F = 16.8$, $p = 5.2e{-}5$), BA35 thickness ($F = 10.4$, $p = 1.4e{-}3$), ERC volume ($F = 9.5$, $p = 2.2e{-}3$), and BA35 volume ($F = 8.2$, $p = 4.5e{-}3$). Volume and thickness of all the subregions were significantly smaller in patient groups at the late prodromal stage and the differences are bigger in dementia. No significant differences were observed in the preclinical stage. However, there was a trend level ($F = 2.8$, $p = .093$) difference in BA35 thickness between preclinical AD patients and controls.

Overall, the FreeSurfer results were similar to that of ASHS-T1, that is, (a) none of the measures from FreeSurfer showed significant differences in the preclinical AD stage; (b) Bonferroni-corrected significant effects were observed in the early prodromal AD stage (hippocampus volume [$F = 23.0$, $p = 2.0e{-}6$] and PRC volume [$F = 9.8$, $p = 1.9e{-}3$]); and (c) all measurements showed significant differences in late prodromal AD and dementia stages. FreeSurfer ERC and PRC thickness were consistently about 50% thicker than the corresponding measurements (ERC and BA35) by ASHS-T1, which is probably due to the mislabeling of dura as cortex. In addition, the mislabeling of dura seems to introduce instability of FreeSurfer measurements of the MTL cortex in Aβ− controls and early stages of AD (preclinical and early prodromal AD). For example, FreeSurfer ERC volume decreased from Aβ− controls (802.5 mm^3) to preclinical AD (768.2 mm^3) but became slightly higher in early prodromal AD (804.1 mm^3). Also, FreeSurfer volume and thickness measurements were more variable (higher SD) than the corresponding measurements generated by the proposal pipeline.

4 | DISCUSSION

In this article, we present an automatic segmentation pipeline for T1w MRI for measuring MTL subregions accounting for the confound of dura and variable anatomy of the MTL cortex. The cross-validation accuracy of ASHS-T1 relative to manual segmentation was relatively high, with DSC ranging from 0.71 to 0.93. The segmentation accuracy of the ASHS-T1 pipeline is comparable to that of our T2 pipeline (except for ERC for which the accuracy is slightly lower, shown in Table 3). Cross-validation experiments in the PMC-T1 atlas showed that ASHS-T1 can reliably separate dura from gray matter, only mislabeling 6.5% of the dura as gray matter, whereas the FreeSurfer mislabels 62.4% of dura as gray matter, leading to about 50% thicker cortex in ERC and PRC. In the ADNI data set, we showed that the degree of dura mislabeling in FreeSurfer decreases with increasing disease severity, indicating a bias where the cortex is oversegmented to a larger extent in Aβ− controls than in patients. This could potentially lead to an overestimation of group differences in later stages of the disease. Finally, in the ADNI data set, we demonstrated that our pipeline picks up changes in early prodromal AD in the MTL, including in ERC and BA35, which agrees with the known progression of NFT pathology, but also in the posterior hippocampus. Moreover, the volume and thickness loss become more severe and widespread with increasing disease severity.

The ASHS-T1 pipeline has several unique aspects and strengths. First, it provides granular measures of the MTL, including subdivision of the PRC and hippocampus, for T1w MRI. It could therefore be very useful in clinical trials and large-scale studies (e.g., ADNI) in older populations in the interrogation of, for example, AD or age-related effects on the MTL, the anterior and posterior MTL networks and memory processes that differentially depend on these regions. In contrast to most previous methods for T1w MRI, our multiatlas approach for labeling MTL cortical regions takes into account the anatomical variability of the MTL cortex, which is known to influence the locations of borders between MTL cortical regions. The accuracy of the automated segmentations generated by ASHS-T1 compared to the manual segmentations is good with a DSC >0.76, except for BA35, and ICC >0.76, except for ERC and PHC. As far as we know, this is the first validation of automated segmentation of MTL cortices against manual segmentations on T1w MRI in an older population and the only other study performing such a validation on T2w MRI was of this same pipeline. ASHS-T1 demonstrated similar accuracy for the MTL cortices as the T2w-based ASHS, with a slightly lower accuracy for the ERC. Two other studies performed an evaluation of MTL cortices in younger adults and patients with temporal lobe epilepsy and reported DSC values in the same range (Hu, Coupé, Pruessner, & Collins, 2014; Kim, Caldairou, Bernasconi, & Bernasconi, 2018). With regard to hippocampus, our pipeline performs comparable to state-of-the-art methods (Collins & Pruessner, 2010; Coupé et al., 2011; Leung et al., 2010; Platero & Tobar, 2016; Wang et al., 2013).

In addition, to the best of our knowledge, this is the first automated pipeline that directly labels dura when segmenting MTL subregions in T1w MRI. Experimental results indicate that the dura can be reliably separated from the gray matter (the DSC and ICC are 0.75 and 0.85, respectively), indicating that the portions of the dura that do not merge with the cortex (white arrows in Figure 2) provide sufficient features for automatic segmentation when there

is low or even no contrast between the dura and the cortex. The importance of accounting for the confound of dura in T1w-MRI was shown in the analyses of dura labeling in the FreeSurfer pipeline which indicated that not taking the dura into account can lead to (a) mislabeling of dura as gray matter causing errors in volume or thickness estimations and (b) a bias where this mislabeling is larger in controls than patients. However, it should be noted that while our pipeline explicitly accounts for the dura, it still makes small errors in some subjects (16 out of 663 ADNI subjects, Section 2.4.3) where small portions of the dura are counted towards the gray matter as shown in Figure S1. With the limited contrast differences between dura and gray matter, this cannot be completely avoided. Moreover, this slight mislabeling of dura may explain the slightly lower accuracy of the ERC, as it is adjacent to the dura for a relatively larger extent than other MTL cortices. With the lower resolution and limited contrast in T1w MRI, it is more difficult to resolve this boundary than in T2w MRI. For that reason, MTL cortex segmentation on high-resolution T2-weighted images is still preferred.

A limitation of ASHS-T1 is that the most anterior portions of the ERC and PRC are not included in the segmentation, which is especially of interest for diseases such as semantic variant primary progressive aphasia which show a clear anterior-to-posterior gradient of atrophy in the MTL (Chan et al., 2001; Davies, Halliday, Xuereb, Kril, & Hodges, 2009). We will include these regions in future work. Moreover, the anterior and posterior border of ERC and PRC are directly determined by the extent of the hippocampal head. This could potentially introduce an error where ERC and PRC volume changes along with hippocampal head volume, that is, if the hippocampal head extends for a lower number of slices, ERC and PRC will automatically do so as well. This will affect volume measures, but likely less so for thickness. Indeed, in later stages of the disease the percentage volume loss in ERC and PRC is larger than the percentage thickness loss, which may reflect this bias. This is not reflected in the p values probably due to larger variance of volume measurements as they are noisier. Qualitatively, a stronger association of anterior hippocampal volume with ERC and BA35 volume is found than with ERC and BA35 thickness, separately for both hemispheres (Left: ERC volume: $\rho = 0.74$, thickness: $\rho = 0.57$; BA35 volume: $\rho = 0.68$, and thickness: $\rho = 0.57$—Right: ERC volume: $\rho = 0.75$, thickness: $\rho = 0.53$, BA35 volume: $\rho = 0.61$, and thickness: $\rho = 0.49$).

In light of above described strengths and limitations, there are certain guidelines that should be followed when using ASHS-T1. Careful assessment of the MRI scans and segmentations is important, with common segmentation errors involving minor mislabeling of the dura and infrequent mislabeling of the lateral aspect of the hippocampus, which was observed in a small number of ADNI subjects. Because of the composition of the PMC-T1 atlas set, the most appropriate target population is older adults and MCI patients. However, we also applied the atlas to images of patients with early AD dementia, and careful quality assessment indicated that the atlas performed well in this population. This matches our recent findings that varying the composition of an atlas set between only controls, only MCI and/or AD patients, or a mixture of the two groups, did not significantly affect segmentation accuracy (Xie et al., 2018). However, care is warranted when this atlas and pipeline are applied to other populations including other ages and diseases, or very different imaging protocols. When this atlas is applied to images acquired at a different platform or with a

different MRI protocol, it is recommended to use the “Heur” output (Step 5 in Section 2.4.2).

To assess the clinical validity and utility of our pipeline, we applied it to the ADNI data set and compared different stages of AD with Aβ− controls on MTL subregional volume and thickness. Compared to the Aβ− controls, we observed a trend difference in BA35 thickness in preclinical AD (Aβ+ controls), a significant difference in ERC volume, BA35 volume and thickness and posterior and total hippocampal volume in early prodromal AD and in all regions in late prodromal AD and dementia. The observed earliest effect on BA35 is consistent with the earliest accumulation of NFT pathology in this region (Braak & Braak, 1995, 1991; Ding et al., 2009). A recent study in a different, only partially overlapping, subset of ADNI showed a similar, but significant, decrease in BA35 thickness in preclinical AD (Wolk et al., 2017) using T2w MRI. The difference in significance may be due to more reliable segmentation of the MTL cortex because of a better contrast and separation of dura in T2w MRI as compared to T1w MRI. In light of the recently published A/T/N model (Jack et al., 2016), in future work, it will be interesting to further select cases who are also tau-positive and investigate whether these subjects show increased neurodegeneration in BA35.

The spreading of atrophy to adjacent ERC and hippocampus in early prodromal AD also matches the known spreading of NFT pathology (Braak & Braak, 1995, 1991; Ding et al., 2009) and other studies investigating MTL atrophy patterns in the early stages (Killiany et al., 2002; Krumm et al., 2016; Olsen et al., 2017; Stoub, Rogalski, Leurgans, Bennett, & deToledo-Morrell, 2010; Xu et al., 2000; Yushkevich, Pluta, et al., 2015). The volume loss in posterior hippocampus, rather than anterior hippocampus, was surprising, given that pathology starts in BA35, part of the PRC, which is thought to be more strongly connected to the anterior hippocampus, at least in the primate MTL (Aggleton, 2012) [although some inconsistency in the literature exists (Witter, Van Hoesen, & Amaral, 1989)]. One might therefore speculate that the anterior hippocampus could be affected earlier than the posterior hippocampus in AD. Only a few studies investigated atrophy in the anterior and posterior hippocampus in MCI, where one study reported specific atrophy in anterior regions (Martin, Smith, Collins, Schmitt, & Gold, 2010), but another did not (Greene & Killiany, 2012). Moreover, a qualitative inspection of studies using shape analysis of the hippocampus to investigate granular effects of MCI shows inconsistent findings not clearly pointing towards an anterior-to-posterior gradient of atrophy in MCI (Apostolova et al., 2012; Chételat et al., 2008; Qiu et al., 2009). Additionally, tractography studies in primates indicate that the posterior hippocampus is more strongly connected with the PHC which is in turn connected via the cingulum bundle with regions such as the posterior cingulate cortex and precuneus [this has also been supported by fMRI studies (Aggleton, 2012; Das et al., 2014; Mufson & Pandya, 1984; Poppenk, Evensmoen, Moscovitch, & Nadel, 2013)] which have been indicated recently to show the earliest amyloid pathology (Palmqvist et al., 2017). This amyloid pathology, which is likely already present for years by the time subjects reach the early MCI stage, may have indirectly affected posterior hippocampal integrity. Moreover, the posterior hippocampus is part of the posterior MTL network (Ranganath & Ritchey, 2012), which has been found to already show atrophy in early MCI (Das, Mancuso, Olson, Arnold, & Wolk, 2016). Finally, the increasing severity and widespread atrophy of the MTL in late prodromal AD and dementia again matches known spreading of NFT pathology (Braak &

Braak, 1995, 1991; Ding et al., 2009) and other in vivo MTL work (de Flores, La Joie, & Chételat, 2015; Dickerson et al., 2001; Jauhiainen et al., 2009; Stoub et al., 2010).

In general, FreeSurfer performed fairly similar in this data set consisting only of T1w MRI scans in characterizing the MTL atrophy pattern in the different AD stages by finding morphometric changes in PRC and hippocampus in early prodromal AD and increasing atrophy, including ERC, at later stages. The most evident difference in the early stages is a lack of significant ERC volume or thickness loss in early prodromal AD using FreeSurfer. In fact, when looking carefully at the ERC volume measures, a fluctuation can be observed where ERC volume loss is observed in preclinical AD compared to controls but then an increase is observed in early prodromal AD, where ERC volumes again match those in the control group. This may be due to mislabeling dura as ERC which may introduce additional noise. Given that ERC atrophy is expected to be subtle at this stage, and that a bias with regard to the dura mislabeling was observed at later disease stages, the inclusion of dura in the ERC label may lead only to increased measurement error. Surprisingly, even though we observed a bias in FreeSurfer of decreasing mislabeling of dura, this did not lead to larger effect sizes for group differences between late prodromal AD or dementia compared to controls. Perhaps this effect is counterbalanced by some other features of the labels, for example, the effect size may be weakened by the larger anterior extent of ERC and PRC in FreeSurfer which may potentially not show equal neurodegeneration along the full length. An important note is that the ERC and PRC in FreeSurfer do not represent completely the same regions as the ERC and the combined BA35 and BA36 into PRC in ASHS-T1. ERC and PRC in FreeSurfer actually show a 44.7 and 40.8% overlap with BA35 of the manual segmentation in our atlas, and in fact, ERC and PRC in FreeSurfer have about 37% overlap with each other. Having a granular label of BA35 rather than including it in ERC or a larger PRC label is advantageous, especially in the earliest stages of AD where NFT pathology is only thought to affect the transentorhinal cortex, which approximates our BA35 label, and a small portion of the lateral ERC. We did observe BA35 thinning in preclinical AD compared to Aβ− controls with our pipeline, although only at a trend level, which could potentially due to the heterogeneity in disease severity of the preclinical group.

5 | CONCLUSIONS

In conclusion, we presented a reliable automated pipeline for obtaining granular measures of MTL subregions in T1w MRI, explicitly accounting for the confound of the dura. We demonstrated the clinical utility of this approach by showing atrophy of early Braak regions in early prodromal AD which becomes more severe and widespread in later stages. These findings should be replicated in other cohorts. Interesting and important future directions are establishing change in MTL regions over time, as longitudinal atrophy is more closely linked to clinical status and is important for tracking disease progression or as potential marker in clinical trials and establishing the association with, for example, tau-PET uptake to better understand the drivers of neurodegeneration. This pipeline could be particularly useful for investigating tau-PET tracers that show high uptake in the dura. We hope that this publicly available atlas and software including a cloud-based service (https://sites.google.com/view/ashs-dox/home and https://sites.google.com/view/ashs-dox/cloud-

ashs/overview) will serve the scientific community and enable the interrogation of the role of the MTL in aging, dementia, and cognition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health (grant numbers R01-AG056014, R01-AG040271, P30-AG010124, R01-EB017255, and R01-AG055005); and the donors of Alzheimer's disease research, a program of the BrightFocus Foundation (L.E.M.W.) and Foundation Philippe Chatrier (R.d.F.). This research was also supported by the Spanish DPI2017–87743-R grant from the Spain Ministry of Economy, Industry and Competitiveness (J.V.M.). 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 Department of Defense 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.; Cogstate and 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 and 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 Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Funding information

Northern California Institute for Research and Education; Foundation for the National Institutes of Health; Canadian Institutes of Health Research; Transition Therapeutics; Takeda Pharmaceutical Company; Servier; Piramal Imaging; Pfizer Inc.; Novartis Pharmaceuticals Corporation; Neurotrack Technologies; NeuroRx Research; Meso Scale Diagnostics, LLC.; Lundbeck and Merck & Co., Inc.; Lumosity; Johnson & Johnson Pharmaceutical Research & Development LLC.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; IXICO Ltd.; GE Healthcare; Fujirebio; Genentech, Inc.; F. Hoffmann-La Roche Ltd.; EuroImmun; Eli Lilly and Company; Elan Pharmaceuticals, Inc.; Cogstate and Eisai Inc.; CereSpir, Inc.; Bristol-Myers Squibb Company; Biogen; BioClinica, Inc.; Araclon Biotech; Alzheimer's Drug Discovery Foundation; Alzheimer's Association; AbbVie; National Institute of Biomedical Imaging and Bioengineering; National Institute on Aging; Department of Defense ADNI, Grant/Award Number: W81XWH-12-2-0012; Alzheimer's Disease Neuroimaging Initiative, Grant/Award Number: U01 AG024904; Spain Ministry of Economy, Industry and Competitiveness, Grant/Award Number: DPI2017–87743-R; Foundation Philippe Chatrier; BrightFocus Foundation; National Institutes of Health, Grant/Award Numbers: R01-AG055005, R01-EB017255, P30-AG010124, R01-AG040271, R01-AG056014

REFERENCES

Aggleton JP (2012). Multiple anatomical systems embedded within the primate medial temporal lobe: Implications for hippocampal function. Neuroscience and Biobehavioral Reviews, 36, 1579–1596. 10.1016/J.NEUBIOREV.2011.09.005 [PubMed: 21964564]

Apostolova LG, Green AE, Babakchanian S, Hwang KS, Chou Y, Toga AW, & Thompson PM (2012). Hippocampal atrophy and ventricular enlargement in Normal aging, mild cognitive impairment (MCI), and Alzheimer disease. Alzheimer Disease and Associated Disorders, 26, 17–27. 10.1097/wad.0b013e3182163b62 [PubMed: 22343374]

Augustinack JC, Huber KE, Stevens AA, Roy M, Frosch MP, van der Kouwe AJW, … Fischl B (2013). Predicting the location of human perirhinal cortex, Brodmann's area 35, from MRI. NeuroImage, 64, 32–42. 10.1016/j.neuroimage.2012.08.071 [PubMed: 22960087]

Avants BB, Epstein CL, Grossman M, & Gee JC (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Medical Image Analysis, 12, 26–41. 10.1016/j.media.2007.06.004 [PubMed: 17659998]

Bender AR, Keresztes A, Bodammer NC, Shing YL, Werkle-Bergner M, Daugherty AM, … Raz N (2018). Optimization and validation of automated hippocampal subfield segmentation across the lifespan. Human Brain Mapping, 39, 916–931. 10.1002/hbm.23891 [PubMed: 29171108]

Berron D, Vieweg P, Hochkeppler A, Pluta JB, Ding S-L, Maass A, Wisse LEM (2017). A protocol for manual segmentation of medial temporal lobe subregions in 7 tesla MRI. NeuroImage Clinical, 15, 466–482. 10.1016/j.nicl.2017.05.022 [PubMed: 28652965]

Bobinski M, Wegiel J, Tarnawski M, Bobinski M, Reisberg B, de Leon MJ, … Wisniewski HM (1997). Relationships between regional neuronal loss and neurofibrillary changes in the hippocampal formation and duration and severity of Alzheimer disease. Journal of Neuropathology and Experimental Neurology, 56, 414–420. [PubMed: 9100672]

Boccardi M, Bocchetta M, Apostolova LG, Barnes J, Bartzokis G, Corbetta G, … Frisoni GB (2015). Delphi definition of the EADC-ADNI harmonized protocol for hippocampal segmentation on magnetic resonance. Alzheimer's & Dementia, 11, 126–138. 10.1016/J.JALZ.2014.02.009

Boccardi M, Bocchetta M, Morency FC, Collins DL, Nishikawa M, Ganzola R, … EADC-ADNI Working Group on The Harmonized Protocol for Manual Hippocampal Segmentation and for the Alzheimer's Disease Neuroimaging Initiative. (2015). Training labels for hippocampal segmentation based on the EADC-ADNI harmonized hippocampal protocol. Alzheimer's & Dementia, 11, 175–183. 10.1016/j.jalz.2014.12.002

Braak H, & Braak E (1991). Neuropathological stageing of Alzheimer-related changes. Acta Neuropathologica, 82, 239–259. [PubMed: 1759558]

Braak H, & Braak E (1995). Staging of Alzheimer's disease-related neuro-fibrillary changes. Neurobiology of Aging, 16, 271–278. [PubMed: 7566337]

Chan D, Fox NC, Scahill RI, Crum WR, Whitwell JL, Leschziner G, … Rossor MN (2001). Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. Annals of Neurology, 49, 433–442. 10.1002/ana.92 [PubMed: 11310620]

Chételat G, Fouquet M, Kalpouzos G, Denghien I, De la Sayette V, Viader F, … Desgranges B (2008). Three-dimensional surface mapping of hippocampal atrophy progression from MCI to AD and over normal aging as assessed using voxel-based morphometry. Neuropsychologia, 46, 1721–1731. 10.1016/J.NEUROPSYCHOLOGIA.2007.11.037 [PubMed: 18289618]

Collins DL, & Pruessner JC (2010). Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. NeuroImage, 52, 1355–1366. 10.1016/J.NEUROIMAGE.2010.04.193 [PubMed: 20441794]

Coupé P, Manjón JV, Fonov V, Pruessner J, Robles M, & Collins DL (2011). Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. NeuroImage, 54, 940–954. 10.1016/J.NEUROIMAGE.2010.09.018 [PubMed: 20851199]

Das SR, Mancuso L, Olson IR, Arnold SE, & Wolk DA (2016). Short-term memory depends on dissociable medial temporal lobe regions in amnestic mild cognitive impairment. Cerebral Cortex, 26, 2006–2017. 10.1093/cercor/bhv022 [PubMed: 25725042]

Das SR, Pluta J, Mancuso L, Kliot D, Yushkevich PA, & Wolk DA (2014). Anterior and posterior MTL networks in aging and MCI. Neurobiology of Aging, 36, S141–S150.e1. 10.1016/j.neurobiolaging.2014.03.041 [PubMed: 25444600]

Davies RR, Halliday GM, Xuereb JH, Kril JJ, & Hodges JR (2009). The neural basis of semantic memory: Evidence from semantic dementia. Neurobiology of Aging, 30, 2043–2052. 10.1016/J.NEUROBIOLAGING.2008.02.005 [PubMed: 18367294]

de Flores R, La Joie R, & Chételat G (2015). Structural imaging of hippocampal subfields in healthy aging and Alzheimer's disease. Neuroscience, 309, 29–50. 10.1016/j.neuroscience.2015.08.033 [PubMed: 26306871]

de Flores R, La Joie R, Landeau B, Perrotin A, Mézenge F, de LaSayette V, … Chételat G (2015). Effects of age and Alzheimer's disease on hippocampal subfields: Comparison between manual and freesurfer volumetry. Human Brain Mapping, 36, 463–474. 10.1002/hbm.22640 [PubMed: 25231681]

De Vita E, Thomas DL, Roberts S, Parkes HG, Turner R, Kinchesh P, … Ordidge RJ (2003). High resolution MRI of the brain at 4.7 tesla using fast spin echo imaging. The British Journal of Radiology, 76, 631–637. 10.1259/bjr/69317841 [PubMed: 14500278]

Delli Pizzi S, Franciotti R, Bubbico G, Thomas A, Onofrj M, & Bonanni L (2016). Atrophy of hippocampal subfields and adjacent extrahippocampal structures in dementia with Lewy bodies and Alzheimer's disease. Neurobiology of Aging, 40, 103–109. 10.1016/J.NEUROBIOLAGING.2016.01.010 [PubMed: 26973109]

Dice LR (1945). Measures of the amount of ecologic association between species. Ecology, 26, 297–302. 10.2307/1932409

Dickerson BC, Goncharova I, Sullivan MP, Forchetti C, Wilson RS, Bennett DA, … deToledo-Morrell L (2001). MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. Neurobiology of Aging, 22, 747–754. 10.1016/S0197-4580(01)00271-8 [PubMed: 11705634]

Ding S-L, & Van Hoesen GW (2010). Borders, extent, and topography of human perirhinal cortex as revealed using multiple modern neuroanatomical and pathological markers. Human Brain Mapping, 31, 1359–1379. 10.1002/hbm.20940 [PubMed: 20082329]

Ding S-L, Van Hoesen GW, Cassell MD, & Poremba A (2009). Parcellation of human temporal polar cortex: A combined analysis of multiple cytoarchitectonic, chemoarchitectonic, and pathological markers. The Journal of Comparative Neurology, 514, 595–623. 10.1002/cne.22053 [PubMed: 19363802]

Ekstrom AD, Bazih AJ, Suthana NA, Al-Hakim R, Ogura K, Zeineh M, … Bookheimer SY (2009). Advances in high-resolution imaging and computational unfolding of the human hippocampus. NeuroImage, 47, 42–49. 10.1016/j.neuroimage.2009.03.017 [PubMed: 19303448]

Fischl B (2012). FreeSurfer. NeuroImage, 62, 774–781. 10.1016/j.neuroimage.2012.01.021 [PubMed: 22248573]

Fischl B, Stevens AA, Rajendran N, Yeo BTT, Greve DN, VanLeemput K, … Augustinack JC (2009). Predicting the location of entorhinal cortex from MRI. NeuroImage, 47, 8–17. 10.1016/j.neuroimage.2009.04.033 [PubMed: 19376238]

Frisoni GB, Jack CR, Bocchetta M, Bauer C, Frederiksen KS, Liu Y, … Winblad B (2015). The EADC-ADNI harmonized protocol for manual hippocampal segmentation on magnetic resonance: Evidence of validity. Alzheimer's & Dementia, 11, 111–125. 10.1016/J.JALZ.2014.05.1756

Fukutani Y, Kobayashi K, Nakamura I, Watanabe K, Isaki K, & Cairns NJ (1995). Neurons, intracellular and extracellular neurofibrillary tangles in subdivisions of the hippocampal cortex in normal ageing and Alzheimer's disease. Neuroscience Letters, 200, 57–60. 10.1016/0304-3940(95)12083-G [PubMed: 8584267]

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, … Jenkinson M (2013). The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage, 80, 105–124. 10.1016/j.neuroimage.2013.04.127 [PubMed: 23668970]

Greene SJ, & Killiany RJ (2012). Hippocampal subregions are differentially affected in the progression to Alzheimer's disease. The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology, 295, 132–140. 10.1002/ar.21493

Hu S, Coupé P, Pruessner JC, & Collins DL (2014). Nonlocal regularization for active appearance model: Application to medial temporal lobe segmentation. Human Brain Mapping, 35, 377–395. 10.1002/hbm.22183 [PubMed: 22987811]

Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, … Van Leemput K (2015). 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. 10.1016/j.neuroimage2015.04.042 [PubMed: 25936807]

Jack CR, Bennett DA, Blennow K, Carrillo MC, Feldman HH, Frisoni GB, … Dubois B (2016). A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. Neurology, 87, 539–547. 10.1212/WNL.0000000000002923 [PubMed: 27371494]

Jack CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, … Weiner MW (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging, 27, 685–691. 10.1002/jmri.21049 [PubMed: 18302232]

Jauhiainen AM, Pihlajamäki M, Tervo S, Niskanen E, Tanila H, Hänninen T, … Soininen H (2009). Discriminating accuracy of medial temporal lobe volumetry and fMRI in mild cognitive impairment. Hippocampus, 19, 166–175. 10.1002/hipo.20494 [PubMed: 18777563]

Killiany RJ, Hyman BT, Gomez-Isla T, Moss MB, Kikinis R, Jolesz F, … Albert MS (2002). MRI measures of entorhinal cortex vs hippocampus in preclinical AD. Neurology, 58, 1188–1196. [PubMed: 11971085]

Kim H, Caldairou B, Bernasconi A, & Bernasconi N (2018). Multi-template mesiotemporal lobe segmentation: Effects of surface and volume feature modeling. Frontiers in Neuroinformatics, 12, 39 10.3389/fninf.2018.00039 [PubMed: 30050423]

Kivisaari S, Probst A, & Taylor K (2013). The perirhinal, entorhinal, and parahippocampal cortices and hippocampus: An overview of functional anatomy and protocol for their segmentation in MR images. fMRI, 239–267.

Krumm S, Kivisaari SL, Probst A, Monsch AU, Reinhardt J, Ulmer S, … Taylor KI (2016). Cortical thinning of parahippocampal subregions in very early Alzheimer's disease. Neurobiology of Aging, 38, 188–196. 10.1016/J.NEUROBIOLAGING.2015.11.001 [PubMed: 26827657]

Landau SM, Mintun MA, Joshi AD, Koeppe RA, Petersen RC, Aisen PS, … Jagust WJ (2012). Amyloid deposition, hypometabolism, and longitudinal cognitive decline. Annals of Neurology, 72, 578–586. 10.1002/ana.23650 [PubMed: 23109153]

Lehmann M, Douiri A, Kim LG, Modat M, Chan D, Ourselin S, …Fox NC (2010). Atrophy patterns in Alzheimer's disease and semantic dementia: A comparison of FreeSurfer and manual volumetric measurements. NeuroImage, 49, 2264–2274. 10.1016/J.NEUROIMAGE.2009.10.056 [PubMed: 19874902]

Leow AD, Klunder AD, Jack CR, Toga AW, Dale AM, Bernstein MA, … Thompson PM (2006). Longitudinal stability of MRI for mapping brain change using tensor-based morphometry. NeuroImage, 31, 627–640. 10.1016/j.neuroimage.2005.12.013 [PubMed: 16480900]

Leung KK, Barnes J, Ridgway GR, Bartlett JW, Clarkson MJ, Macdonald K, … Alzheimer's Disease Neuroimaging Initiative, ADN. (2010). Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. NeuroImage, 51, 1345–1359. 10.1016/j.neuroimage.2010.03.018 [PubMed: 20230901]

Mah L, Binns MA, & Steffens DC (2015). Anxiety symptoms in amnestic mild cognitive impairment are associated with medial temporal atrophy and predict conversion to Alzheimer disease. The American Journal of Geriatric Psychiatry, 23, 466–476. 10.1016/J.JAGP.2014.10.005 [PubMed: 25500120]

Malykhin NV, Bouchard TP, Camicioli R, & Coupland NJ (2008). Aging hippocampus and amygdala. Neuroreport, 19, 543–547. 10.1097/wnr.0b013e3282f8b18c [PubMed: 18388735]

Malykhin NV, Bouchard TP, Ogilvie CJ, Coupland NJ, Seres P, & Camicioli R (2007). Three-dimensional volumetric analysis and reconstruction of amygdala and hippocampal head, body and tail. Psychiatry Research: Neuroimaging, 155, 155–165. 10.1016/J.PSCYCHRESNS.2006.11.011

Manjón JV, Coupé P, Buades A, Fonov V, Collins LD, & Robles M (2010). Non-local MRI upsampling. Medical Image Analysis, 14, 784–792. 10.1016/j.media.2010.05.010 [PubMed: 20566298]

Martin SB, Smith CD, Collins HR, Schmitt FA, & Gold BT (2010). Evidence that volume of anterior medial temporal lobe is reduced in seniors destined for mild cognitive impairment. Neurobiology of Aging, 31, 1099–1106. 10.1016/J.NEUROBIOLAGING.2008.08.010 [PubMed: 18809227]

Mishra S, Gordon BA, Su Y, Christensen J, Friedrichsen K, Jackson K, … Benzinger TLS (2017). AV-1451 PET imaging of tau pathology in preclinical Alzheimer disease: Defining a summary measure. NeuroImage, 161, 171–178. 10.1016/J.NEUROIMAGE.2017.07.050 [PubMed: 28756238]

Mueller SG, Stables L, Du AT, Schuff N, Truran D, Cashdollar N, & Weiner MW (2007). Measurement of hippocampal subfields and age-related changes with high resolution MRI at 4T. Neurobiology of Aging, 28, 719–726. 10.1016/j.neurobiolaging.2006.03.007 [PubMed: 16713659]

Mufson EJ, & Pandya DN (1984). Some observations on the course and composition of the cingulum bundle in the rhesus monkey. The Journal of Comparative Neurology, 225, 31–43. 10.1002/cne.902250105 [PubMed: 6725639]

Olsen RK, Palombo DJ, Rabin JS, Levine B, Ryan JD, & Rosenbaum RS (2013). Volumetric analysis of medial temporal lobe subregions in developmental amnesia using high-resolution magnetic resonance imaging. Hippocampus, 23, 855–860. 10.1002/hipo.22153 [PubMed: 23749334]

Olsen RK, Yeung L-K, Noly-Gandon A, D'Angelo MC, Kacollja A, Smith VM, … Barense MD (2017). Human anterolateral entorhinal cortex volumes are associated with cognitive decline in aging prior to clinical diagnosis. Neurobiology of Aging, 57, 195–205. 10.1016/J.NEUROBIOLAGING.2017.04.025 [PubMed: 28578804]

Palmqvist S, Schöll M, Strandberg O, Mattsson N, Stomrud E, Zetterberg H, … Hansson O (2017). Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity. Nature Communications, 8, 1214 10.1038/s41467-017-01150-x

Pasquini L, Scherr M, Tahmasian M, Myers NE, Ortner M, Kurz A,… Sorg C (2016). Increased intrinsic activity of medial-temporal lobe subregions is associated with decreased cortical thickness of medialparietal areas in patients with Alzheimer's disease dementia. Journal of Alzheimer's Disease, 51, 313–326. 10.3233/JAD-150823

Petersen RC (2004). Mild cognitive impairment as a diagnostic entity. Journal of Internal Medicine, 256, 183–194. 10.1111/j.1365-2796.2004.01388.x [PubMed: 15324362]

Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, … Jack CR (2009). Mild cognitive impairment: Ten years later. Archives of Neurology, 66, 1447–1455. 10.1001/archneurol.2009.266 [PubMed: 20008648]

Platero C, & Tobar MC (2016). A fast approach for hippocampal segmentation from T1-MRI for predicting progression in Alzheimer's disease from elderly controls. Journal of Neuroscience Methods, 270, 61–75. 10.1016/j.jneumeth.2016.06.013 [PubMed: 27328371]

Poppenk J, Evensmoen HR, Moscovitch M, & Nadel L (2013). Long-axis specialization of the human hippocampus. Trends in Cognitive Sciences, 17, 230–240. 10.1016/j.tics.2013.03.005 [PubMed: 23597720]

Preston AR, Bornstein AM, Hutchinson JB, Gaare ME, Glover GH, & Wagner AD (2010). High-resolution fMRI of content-sensitive subsequent memory responses in human medial temporal lobe. Journal of Cognitive Neuroscience, 22, 156–173. 10.1162/jocn.2009.21195 [PubMed: 19199423]

Qiu A, Fennema-Notestine C, Dale AM, Miller MI, & Alzheimer's Disease Neuroimaging Initiative, ADN. (2009). Regional shape abnormalities in mild cognitive impairment and Alzheimer's disease. NeuroImage, 45, 656–661. [PubMed: 19280688]

Ranganath C, & Ritchey M (2012). Two cortical systems for memory-guided behaviour. Nature Reviews. Neuroscience, 13, 713–726. 10.1038/nrn3338 [PubMed: 22992647]

Shrout PE, & Fleiss JL (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86, 420–428. [PubMed: 18839484]

Stoub TR, Rogalski EJ, Leurgans S, Bennett DA, & deToledo-Morrell L (2010). Rate of entorhinal and hippocampal atrophy in incipient and mild AD: Relation to memory function. Neurobiology of Aging, 31, 1089–1098. 10.1016/J.NEUROBIOLAGING.2008.08.003 [PubMed: 18809228]

Thomas DL, De Vita E, Roberts S, Turner R, Yousry TA, & Ordidge RJ (2004). High-resolution fast spin echo imaging of the human brain at 4.7 T: Implementation and sequence characteristics. Magnetic Resonance in Medicine, 51, 1254–1264. 10.1002/mrm.20106 [PubMed: 15170847]

Wang H, Pouch A, Takabe M, Jackson B, Gorman J, Gorman R, & Yushkevich PA (2013). Multi-atlas segmentation with robust label transfer and label fusion. Information Processing in Medical Imaging, 23, 548–559. 10.1007/978-3-642-38868-2_46 [PubMed: 24683998]

Wang H, & Yushkevich PA (2013). Multi-atlas segmentation with joint label fusion and corrective learning-an open source implementation. Frontiers in Neuroinformatics, 7, 27 10.3389/fninf.2013.00027 [PubMed: 24319427]

Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund L-O, … Petersen RC (2004). Mild cognitive impairment--beyond controversies, towards a consensus: Report of the international working group on mild cognitive impairment. Journal of Internal Medicine, 256, 240–246. 10.1111/j.1365-2796.2004.01380.x [PubMed: 15324367]

Wisse LEM, Gerritsen L, Zwanenburg JJM, Kuijf HJ, Luijten PR, Biessels GJ, & Geerlings MI (2012). Subfields of the hippocampal formation at 7T MRI: in vivo volumetric assessment. NeuroImage, 61, 1043–1049. 10.1016/j.neuroimage.2012.03.023 [PubMed: 22440643]

Wisse LE, Biessels GJ, & Geerlings MI (2014). A critical appraisal of the hippocampal subfield segmentation package in FreeSurfer. Frontiers in aging neuroscience, 6, 261. [PubMed: 25309437]

Witter MP, Van Hoesen GW, & Amaral DG (1989). Topographical organization of the entorhinal projection to the dentate gyrus of the monkey. The Journal of Neuroscience, 9, 216–228. [PubMed: 2913203]

Wolk DA, Das SR, Mueller SG, Weiner MW, Yushkevich PA, & Alzheimer's Disease Neuroimaging Initiative. (2017). Medial temporal lobe subregional morphometry using high resolution MRI in Alzheimer's disease. Neurobiology of Aging, 49, 204–213. 10.1016/j.neurobiolaging.2016.09.011 [PubMed: 27836336]

Wolk DA, & Dickerson BC (2011). Fractionating verbal episodic memory in Alzheimer's disease. NeuroImage, 54, 1530–1539. 10.1016/J.NEUROIMAGE.2010.09.005 [PubMed: 20832485]

Xie L, Pluta J, Wang H, Das SR, Mancuso L, Kliot D, … Yushkevich PA (2014). Automatic clustering and thickness measurement of anatomical variants of the human perirhinal cortex. Medical Image Computing and Computer-Assisted Intervention–MICCAI, 2014, 81–88.

Xie L, Pluta JB, Das SR, Wisse LEM, Wang H, Mancuso L, … Yushkevich PA (2017). Multi-template analysis of human perirhinal cortex in brain MRI: Explicitly accounting for anatomical variability. NeuroImage, 144, 183–202. 10.1016/j.neuroimage.2016.09.070 [PubMed: 27702610]

Xie L, Shinohara RT, Ittyerah R, Kuijf HJ, Pluta JB, Blom K, … Wisse LEM, (2018). Automated multi-atlas segmentation of hippocampal and extrahippocampal subregions in Alzheimer's disease at 3T and 7T: What atlas composition works best? Journal of Alzheimer's Disease, 63, 217–225. 10.3233/JAD-170932

Xie L, Wisse LEM, Das SR, Wang H, Wolk DA, Manjón JV, & Yushkevich PA (2016). Accounting for the confound of meninges in segmenting entorhinal and perirhinal cortices in T1-weighted MRI. Medical Image Computing and Computer-Assisted Intervention-MICCAI, 2016, 564–571. 10.1007/978-3-319-46723-8_65

Xu Y, Jack CR, O'Brien PC, Kokmen E, Smith GE, Ivnik RJ, … Petersen RC (2000). Usefulness of MRI measures of entorhinal cortex versus hippocampus in AD. Neurology, 54, 1760–1767. [PubMed: 10802781]

Yonelinas AP, Widaman K, Mungas D, Reed B, Weiner MW, & Chui HC (2007). Memory in the aging brain: Doubly dissociating the contribution of the hippocampus and entorhinal cortex. Hippocampus, 17, 1134–1140. 10.1002/hipo.20341 [PubMed: 17636547]

Yushkevich PA, Amaral RSC, Augustinack JC, Bender AR, Bernstein JD, Boccardi M, … Zeineh MM (2015). Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: Towards a harmonized segmentation protocol. NeuroImage, 111, 526–541. 10.1016/j.neuroimage.2015.01.004 [PubMed: 25596463]

Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, & Gerig G (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. NeuroImage, 31, 1116–1128. 10.1016/j.neuroimage.2006.01.015 [PubMed: 16545965]

Yushkevich PA, Pluta JB, Wang H, Xie L, Ding S, Gertje EC, … Wolk DA (2015). Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. Human Brain Mapping, 36, 258–287. 10.1002/hbm.22627 [PubMed: 25181316]

Zeineh MM, Engel SA, Thompson PM, & Bookheimer SY (2001). Unfolding the human hippocampus with high resolution structural and functional MRI. The Anatomical Record, 265, 111–120. [PubMed: 11323773]

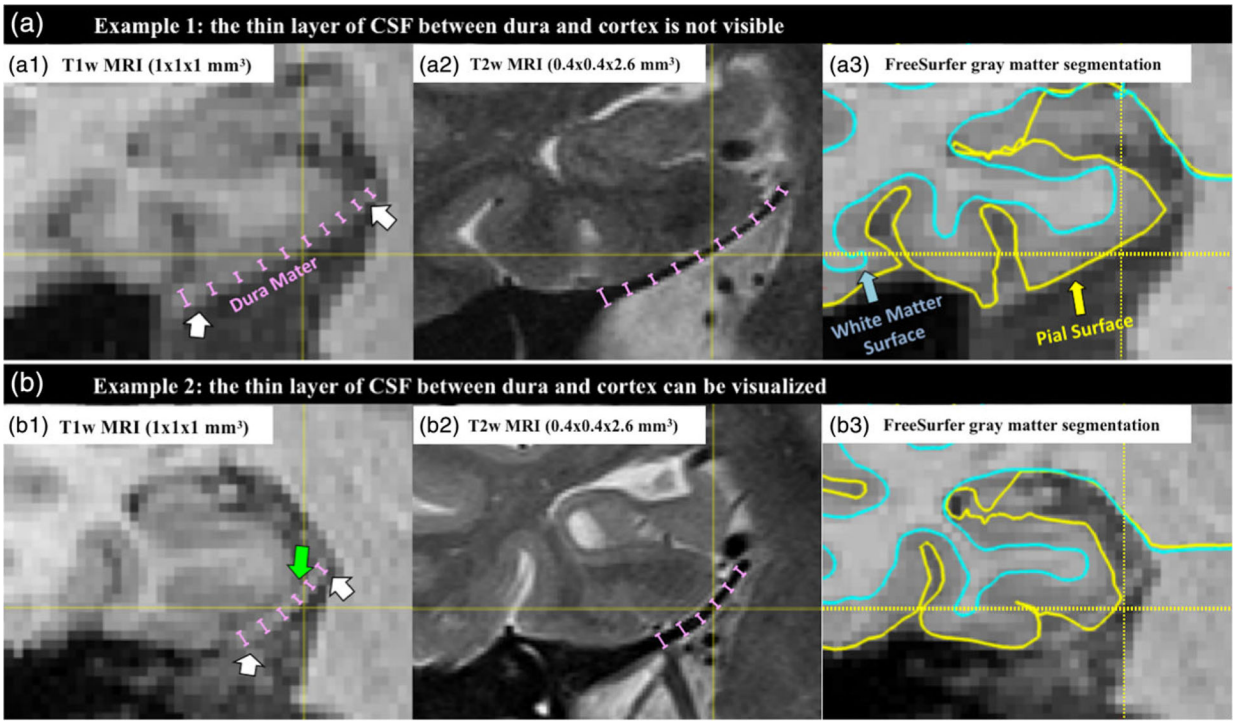

FIGURE 1.

The dura mater (indicated by purple lines) has similar intensity as gray matter in T1w MRI (the first column) but can be easily separated from the cortex in T2w MRI (the hypointensity layer in the second column). It is often segmented as part of the cortex by state-of-the-art algorithms, for example, FreeSurfer (the third column). A thin layer of CSF (green arrow) can be visualized in some subjects (second row, b) but not in the others (first row, a). White arrows point to the portions of the dura that are not merged with the cortex. CSF, cerebrospinal fluid; MRI, magnetic resonance imaging; T1w MRI, T1-weighted MRI; T2w MRI, T2-weighted MRI [Color figure can be viewed at wileyonlinelibrary.com]

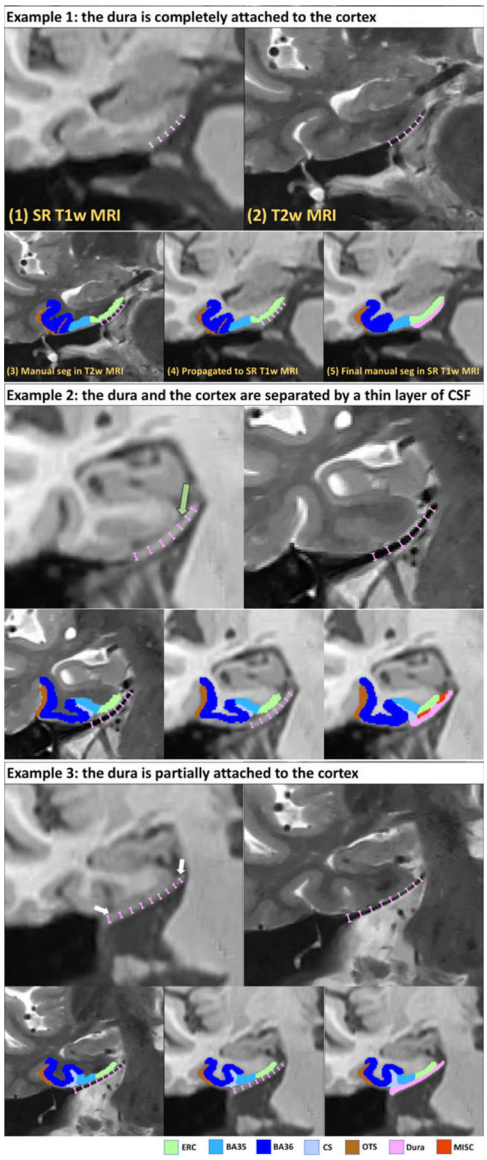

FIGURE 2.

Examples showing the procedure of manual segmenting MTL cortices in T1w MRI using manual labels in the space of the T2w MRI. The dura mater is shown as the thin hypointensity layer in T2w MRI, which appears gray in T1w MRI, indicated by the purple lines. The green arrow points to a thin layer of CSF between dura and the cortex that exists in some subjects. The white arrows show places that the dura are not completely attached to the cortex, which are important features for manual and automatic segmentation. CSF, cerebrospinal fluid; MRI, magnetic resonance imaging; MTL, medial temporal lobe; T1w MRI, T1-weighted MRI; T2w MRI, T2-weighted MRI [Color figure can be viewed at wileyonlinelibrary.com]

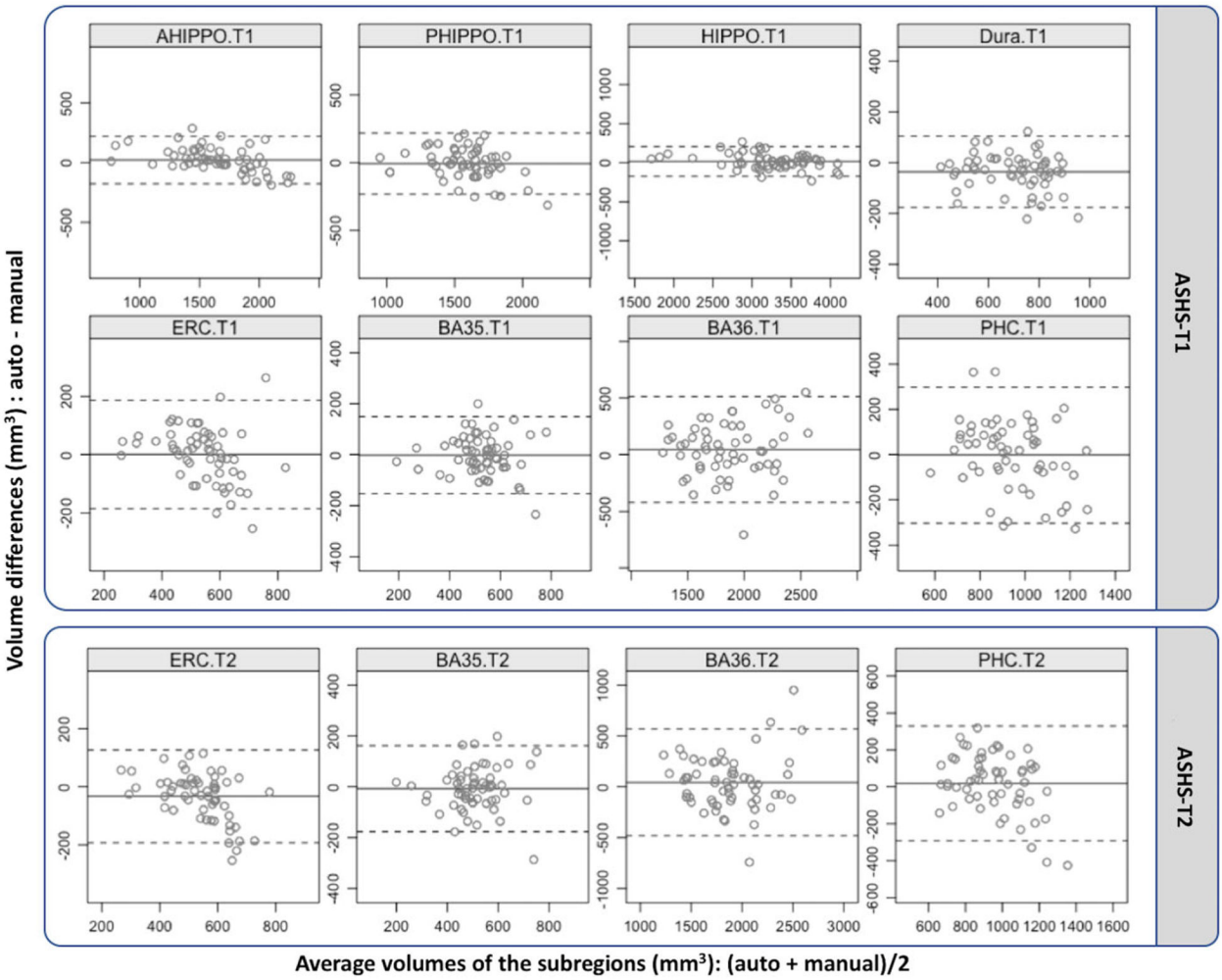

FIGURE 3.

Comparison of subregion volume measured by ASHS-T1 and manual segmentation in the space of T1w MRI (top panel) MRI using Bland–Altman plots. To compare with ASHS-T2 (bottom panel), the comparisons between subregion volume generated in the space of T2w MRI and the corresponding manual segmentation in the PMC-T2 atlas set are also shown (.T2). HIPPO.T1 is a compound label by merging AHIPPO.T1 and PHIPPO.T1. AHIPPO, anterior hippocampus; ASHS, automatic segmentation of hippocampal subfields software; BA35/36, Brodmann area 35/36; ERC, entorhinal cortex; HIPPO, whole hippocampus; MRI, magnetic resonance imaging; PHC, parahippocampal cortex; PHIPPO, posterior hippocampus; PMC, Penn Memory Center; T1w MRI, T1-weighted MRI; T2w MRI, T2-weighted MRI [Color figure can be viewed at wileyonlinelibrary.com]

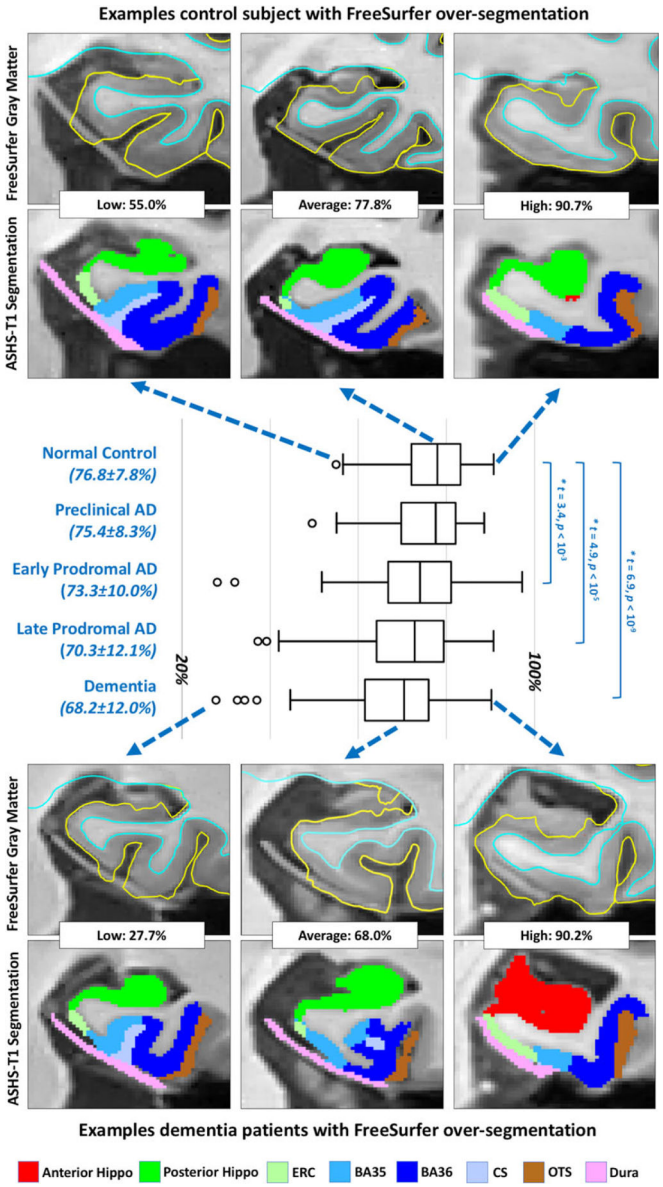

FIGURE 4.

Percentage of dura voxels labeled as gray matter by FreeSurfer in all diagnosis groups. Examples of low, average, and high oversegmentation in controls and dementia patients are shown on top and bottom, respectively. AD, Alzheimer's disease [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1

Characteristics of the PMC atlas set

	Normal control	aMCI
N	15	14
Age (years)	66.3 (9.5)	71.9 (6.2)
Gender (M/F)	7/8	6/8
Education (years)	15.6 (2.6)	16.9 (2.8)
MMSE	29.5 (1.0)	26.9 (1.7)***

Note. All statistics are in comparison to cognitive normal control subjects. Independent two-sample *t* test (age), Wilcoxon rank sum test (education, MMSE), and contingency χ^2 test (gender) were used. *SD* in parentheses.

Abbreviations: aMCI, amnestic mild cognitive impairment; MMSE, mini-mental state examination; PMC, Penn Memory Center.

*** $p < .001$.

TABLE 2

Characteristics of the ADNI data set

	Aβ— control	Preclinical AD	Early prodromal AD	Late prodromal AD	Dementia
N	190	95	142	109	127
Age (years)	72.3 (6.0)	74.8 (5.9)***	73.6 (6.9)	72.3 (6.8)	74.4 (8.2)*
Gender (M/F)	100/90	31/64**	81/61	57/52	68/59
Education (years)	16.9 (2.4)	16.1 (2.7)*	15.6 (2.8)***	16.6 (2.6)	15.6 (2.7)***
MMSE	29.0 (1.3)	29.0 (1.1)	28.0 (1.7)***	27.2 (1.9)***	23.0 (2.1)***

Note. All statistics are in comparison to amyloid-β negative (Aβ−) control subjects. Independent two-sample *t* test (age), Wilcoxon rank sum test (education, MMSE), and contingency χ^2 test (gender) were used. *SD* in parentheses.

Abbreviations: Aβ, beta-amyloid; AD, Alzheimer's disease; MMSE, mini-mental state examination.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

TABLE 3

Segmentation accuracy (measured by DSC) and ICC relative to manual segmentations using leave-one-out cross validation. The atlas consists of 14 aMCI patients and 15 cognitively NC. Volume measurements of each subregion of the manual segmentations are also included in the table. DSCs and volume of each subregion in both hemispheres are averaged. Mean and *SD* (parentheses) are reported in the table

	T1w MRI			T2w MRI		
Subregion	**Volume (mm^3)**	**ICC**	**DSC**	**Volume (mm^3)**	**ICC**	**DSC**
Anterior hippocampus	1,635.3 (363.6)	0.95	0.92 (0.02)	\	\	\
Posterior hippocampus	1,585.8 (258.1)	0.89	0.90 (0.02)	\	\	\
Whole hippocampus[a]	3,221.1 (535.4)	0.98	0.93 (0.01)	\	\	\
ERC	542.0 (258.1)	0.69	0.76 (0.03)	555.6 (130.2)	0.71	0.79 (0.03)
BA35	527.2 (117.6)	0.77	0.71 (0.06)	511.9(111.9)	0.71	0.71 (0.06)
BA36	1842.9 (333.7)	0.76	0.79 (0.03)	1852.5 (349.3)	0.72	0.79 (0.04)
PHC	941.9 (200.6)	0.64	0.80 (0.03)	946.9 (214.3)	0.64	0.79 (0.04)
Dura mater	717.4 (148.3)	0.85	0.75 (0.05)	\	\	\

Abbreviations: aMCI, amnestic mild cognitive impairment; BA35, Brodmann area 35; BA36, Brodmann area 36; DSC, Dice similarity coefficient; ERC, entorhinal cortex; ICC, intraclass correlation coefficient; MRI, magnetic resonance imaging; NC, normal control; PHC, parahippocampal cortex; T1w, T1-weighted; T2w, T2-weighted.

[a]DSC and volume of the compound labels (in italics) are measured using the merged label of corresponding sublabels (whole hippocampus: anterior and posterior hippocampus in T1-weighted MRI).

TABLE 4

Comparisons of different analysis methods in labeling the dura mater in the PMC atlas set

	% of dura voxels in manual segmentation labeled as		
Method	**Dura**	**Gray matter**	**Background and CSF**
ASHS-T1	71.9 ± 6.4	6.5 ± 3.7	21.6 ± 5.9
FreeSurfer 6.0	N/A	62.4 ± 10.5	37.6 ± 10.5

Abbreviations: ASHS-T1, Automatic Segmentation of Hippocampal Subfields; CSF, cerebrospinal fluid; PMC, Penn Memory Center.

TABLE 5

Statistical analysis results using volumetric measurements, adjusted for age and intracranial volume, in discriminating patient groups from Aβ−controls in ADNI. Measurements that survived Bonferroni correction ($p < .05/10$) are highlighted in bold

Region	Aβ–control (*n* = 190)	Preclinical AD (*n* = 95)	Early prodromal AD (*n* = 142)	Late prodromal AD (*n* = 109)	Dementia (*n* = 127)
ASHS–T1 volume measurements (mm^3)					
Anterior hippocampus	1,724.5	1,711.6	1,666.7	1,533.9	1,440.0
SD	225.6	207.0	237.9	255.5	200.9
%Diff		−0.7	−3.3	−11.1	−16.5
F stats		<2.5	4.9	44.8	129.8
p value		>0.1	0.027	l.le−10	2.2e−25
Cohen's d		<0.1	0.25	0.79	1.3
Posterior hippocampus	1,646.5	1,647.0	1,571.4	1,405.0	1,344.1
SD	160.1	152.7	172.0	202.7	167.0
% Diff		0.0	−4.6	−14.7	221218.4
F stats		<2.5	16.8	129.3	261.0
p value		>0.1	5.2e−5	4.4e−25	4.0e−43
Cohen's d		<0.1	0.45	1.32	1.85
Hippocampus	3,371.0	3,358.6	3,238.2	2,938.9	2,784.0
SD	309.9	298.7	351.5	411.8	320.8
% Diff		−0.4	−3.9	−12.8	−17.4
F stats		<2.5	13.1	105.3	262.1
p value		>0.1	3.4e−4	2.5e−21	3.0e−43
Cohen's d		<0.1	0.40	1.19	1.86
ERC	573.9	562.5	548.0	499.6	451.4
SD	73.8	69.8	80.5	91.8	84.6
% Diff		−5.2	−6.1	−12.8	−23.7
F stats		<2.5	9.5	58.5	192.5
p value		>0.1	2.2e−3	2.9e−13	1.9e−34
Cohen's d		0.16	0.34	0.89	1.54
BA35	606.5	597.1	579.6	539.8	481.3
SD	82.2	80.3	92.1	100.2	81.3

Region	Aβ–control (*n* = 190)	Preclinical AD (*n* = 95)	Early prodromal AD (*n* = 142)	Late prodromal AD (*n* = 109)	Dementia (*n* = 127)
% Diff		−1.6	−4.4	−11.0	−20.6
F stats		<2.5	8.2	38.6	183.3
p value		>0.1	4.5e−3	1.8e−9	3.5e−33
Cohen's d		0.12	0.31	0.73	1.53
BA36	1,881.0	1,869.5	1,814.6	1,718.6	1,581.9
SD	249.7	228.5	227.9	264.3	230.3
96 Diff		−0.6	−3.5	−8.6	−15.9
F stats		<2.5	6.1	27.9	115.4
p value		>0.1	0.014	2.5e−7	4.0e−23
Cohen's d		<0.1	0.28	0.63	1.25
PHC	958.0	977.6	959.2	909.7	869.4
SD	117.9	129.0	138.3	129.5	122.8
% Diff		2.1	0.1	−5.0	−9.3
F stats		<2.5	<2.5	10.9	43.5
p value		>0.1	>0.1	1.0e−3	1.8e−10
Cohen's d		<0.1	<0.1	0.39	0.74
FreeSurfer volume measurements (mm^3)					
Hippocampus	3,790.2	3,732.6	3,606.2	3,299.6	3,146.8
SD	321.9	364.1	363.8	425.9	330.8
% Diff		−1.5	−4.9	−12.9	−17.0
F stats		<2.5	23.0	127,2	301.5
p value		>0.1	2.0e−6	9.2e−25	9.0e−48
Cohen's d		0.17	0.54	1.30	1.97
ERC	802.5	768.2	804.1	699.5	662.1
SD	197.1	203.1	212.2	229.8	191.2
% Diff		−4.3	0.2	−12.8	−17.5
F stats		<2.5	<2.5	17.0	38.9
p value		>0.1	>0.1	4.9e−5	1.5e−9
Cohen's d		0.17	<0.1	0.48	0.72
PRC	1,067.5	1,051.9	1,019.0	916.8	834.1
SD	150.0	142.1	147.9	193.4	161.6

Region	Aβ–control (*n* = 190)	Preclinical AD (*n* = 95)	Early prodromal AD (*n* = 142)	Late prodromal AD (*n* = 109)	Dementia (*n* = 127)
% Diff		−1.5	−4.5	−14.1	−21.9
F stats		<2.5	9.8	56.9	171.6
p value		>0.1	1.9e−3	5.7e−13	1.5e−31
Cohen's d		0.11	0.33	0.87	1.50

Abbreviations: Aβ, beta-amyloid; AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; ASHS, Automated Segmentation of Hippocampal Subfields; BA, Brodmann area, ERC, entorhinal cortex; PHC, parahippocampal cortex; PRC, perirhinal cortex.

TABLE 6

Statistical analysis results using thickness measurements, adjusted for age, in discriminating patient groups from Aβ− controls in ADNI. Measurements that survived Bonferroni correction ($p < .05/6$) are highlighted in bold

Region	Aβ–control (*n* = 190)	Preclinical AD (*n* = 95)	Early prodromal AD (*n* = 142)	Late prodromal AD (*n* = 109)	Dementia (*n* = 127)
ASHS−T1 thickness measurements (mm)					
ERC	2.02	2.02	1.99	1.93	1.78
SD	0.16	0.17	0.18	0.18	0.22
% Diff		0.1	−1.2	−4.3	−11.5
F stats		<2.5	<2.5	18.4	122.0
p value		>0.1	>0.1	2.4e−5	3.5e−24
Cohen's *d*		<0.1	0.18	0.53	1.25
BA35	2.35	2.33	2.29	2.20	2.06
SD	0.16	0.18	0.19	0.23	0.23
96 Diff		−1.1	−2.6	−6.7	−12.5
F stats		2.8	10.4	49.1	184.4
p value		0.093	1.4e−3	1.6e−11	2.3e−33
Cohen's *d*		0.12	0.34	0.76	1.46
BA36	2.41	2.39	2.39	2.29	2.21
SD	0.23	0.23	0.21	0.22	0.24
% Diff		−0.7	−0.8	−5.1	−8.5
F stats		<2.5	<2.5	20.3	57.8
p value		>0.1	>0.1	1.0e−5	3.4e−13
Cohen's *d*		<0.1	<0.1	0.53	0.85
PHC	2.15	2.17	2.14	2.09	2.00
SD	0.13	0.16	0.13	0.15	0.16
% Diff		0.7	−0.3	−2.7	−7.2
F stats		<2.5	<2.5	13.3	94.2
p value		>0.1	>0.1	3.2e−4	1.2e−19
Cohen's *d*		<0.1	<0.1	0.43	1.03
FreeSurfer thickness measurements (mm)					
ERC	3.18	3.17	3.14	2.90	2.74
SD	0.30	0.33	0.34	0.40	0.40

Region	Aβ–control (*n* = 190)	Preclinical AD (*n* = 95)	Early prodromal AD (*n* = 142)	Late prodromal AD (*n* = 109)	Dementia (*n* = 127)
% Diff		−0.2	−1.1	−8.8	−13.6
F stats		<2.5	<2.5	46.3	122.6
p value		>0.1	>0.1	5.7e−11	2.9e−24
Cohen's d		<0.1	0.12	0.79	1.24
PRC	3.44	3.47	3.36	3.11	2.87
SD	0.31	0.33	0.34	0.41	0.37
% Diff		1.0	−2.2	−9.5	−16.4
F stats		<2.5	5.2	59.8	217.0
p value		>0.1	0.023	1.7e−13	1.1e−37
Cohen's *d*		<0.1	0.25	0.91	1.67

ADNI. Measurements that survived Bonferroni correction ($p < .05/6$) are highlighted in bold

Abbreviations: Aβ, beta-amyloid; AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; ASHS, Automated Segmentation of Hippocampal Subfields; BA, Brodmann area, ERC, entorhinal cortex; PHC, parahippocampal cortex; PRC, perirhinal cortex.