ELSEVIER

CrossMark

Alzheimer's & Dementia 11 (2015) 184–194

Alzheimer's & Dementia

Operationalizing protocol differences for EADC-ADNI manual hippocampal segmentation

Marina Boccardi[a], Martina Bocchetta[a,b], Rossana Ganzola[a], Nicolas Robitaille[c], Alberto Redolfi[a], Simon Duchesne[c], Clifford R. Jack, Jr.,[d], Giovanni B. Frisoni[a,e,*], for the EADC-ADNI Working Group on The Harmonized Protocol for Manual Hippocampal Segmentation and for the Alzheimer's Disease Neuroimaging Initiative

[a]*Laboratory of Epidemiology, Neuroimaging and Telemedicine, IRCCS - S. Giovanni di Dio – Fatebenefratelli, Brescia, Italy*
[b]*Associazione Fatebenefratelli per la Ricerca, Rome, Italy*
[c]*Department of Radiology, Université Laval and Centre de Recherche de l'Institut universitaire de santé mentale de Québec, Quebec City, Canada*
[d]*Department of Diagnostic Radiology, Mayo Clinic and Foundation, Rochester, MN, USA*
[e]*University Hospitals and University of Geneva, Geneva, Switzerland*

Abstract

Background: Hippocampal volumetry on magnetic resonance imaging is recognized as an Alzheimer's disease (AD) biomarker, and manual segmentation is the gold standard for measurement. However, a standard procedure is lacking. We operationalize and quantitate landmark differences to help a Delphi panel converge on a set of landmarks.

Methods: One hundred percent of anatomic landmark variability across 12 different protocols for manual segmentation was reduced into four segmentation units (the minimum hippocampus, the alveus/fimbria, the tail, and the subiculum), which were segmented on magnetic resonance images by expert raters to estimate reliability and AD-related atrophy.

Results: Intra- and interrater reliability were more than 0.96 and 0.92, respectively, except for the alveus/fimbria, which were 0.86 and 0.77, respectively. Of all AD-related atrophy, the minimum hippocampus contributed to 67%; tail, 24%; alveus/fimbria, 4%; and the subiculum, 5%.

Conclusions: Anatomic landmark variability in available protocols can be reduced to four discrete and measurable segmentation units. Their quantitative assessment will help a Delphi panel to define a set of landmarks for a harmonized protocol.

Keywords: Hippocampus; Hippocampal atrophy; Hippocampal volumetry; Manual segmentation protocol; Harmonization; Anatomic landmark; Alzheimer's disease; Manual tracing; Medial temporal lobes; Atrophy; Degeneration; Magnetic resonance; Neuroimaging; Alzheimer's Disease Neuroimaging Initiative; Standard operating procedures

Project collaborators include: George Bartzokis (Department of Psychiatry, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA), John G. Csernansky (Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA), Mony J. de Leon (New York University School of Medicine, Center for Brain Health, New York, NY, USA), Leyla deToledo-Morrell (Department of Neurological Sciences, Rush University, Chicago, IL, USA), Ronald J. Killiany (Department of Anatomy and Neurobiology, Boston University School of Medicine, Boston, MA, USA), Stéphane Lehéricy (Center for NeuroImaging Research and Department of Neuroradiology, Université Pierre et Marie Curie-Paris 6, Groupe Hospitalier Pitié-Salpêtrière, Paris, France), Nikolai Malykhin (Department of Biomedical Engineering, Centre for Neuroscience, University of Alberta, Edmonton, Alberta, Canada), Johannes Pantel (Institute of General Practice, University of Frankfurt/Main, Germany), Jens C. Pruessner (McGill Centre for Studies in Aging, Department of Psychiatry, McGill University, Montreal, Quebec, Canada), Hilkka Soininen (Department of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland), and Craig Watson (Wayne State University School of Medicine, University Health Center, St. Antoine, Detroit, MI, USA).

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). 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.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. This manuscript was approved by the ADNI Data and Publication Committee on June 4, 2012.

*Corresponding author. Tel.: +39 030 3501361; Fax: +39 030 3501592.
E-mail address: gfrisoni@fatebenefratelli.it

1552-5260/$ - see front matter
http://dx.doi.org/10.1016/j.jalz.2013.03.001

1. Introduction

Hippocampal volumetry has been proposed as a diagnostic marker for Alzheimer's disease (AD) by the International Working Group for the new research criteria for the diagnosis of AD and National Institute of Aging–Alzheimer's Association revised diagnostic criteria [1–5]. To obtain such measurement, manual segmentation on T1-weighted, high-resolution magnetic resonance (MR) images is regarded as the gold standard. Although many algorithms for automated hippocampal segmentation exist (cf. [6] for a cursory review), with undeniable usefulness (less time-consuming, high test–retest reliability), manual segmentation [7–10] remains the benchmark that should be used for the validation of the algorithms carrying out automated segmentation [6,11–15].

However, different manual segmentation protocols are available [16,17], leading to volume estimates differing up to 2.5-fold for hippocampal volume in control subjects [16]. Heterogeneity in anatomic definitions and tracing guidelines have hampered comparisons among different studies using hippocampal volumetry for diagnosis or as a surrogate marker for disease progression, and limit its use as a diagnostic marker for clinical diagnosis.

An effort has been undertaken by European Alzheimer's Disease Consortium (EADC) and Alzheimer's Disease Neuroimaging Initiative (ADNI) centers to develop a harmonized protocol for the manual segmentation of the hippocampus on MR images [18–20]). This project aims to achieve a consensus on a gold standard for manual segmentation through a Delphi panel of experts, and then validate it versus local protocols and pathology.

To achieve this goal, the following steps were planned and executed. First, a survey of the most cited protocols in the literature was undertaken to provide the exact range of anatomic landmark variability, after certification of the protocols by authors, who checked and corrected the practical execution of segmentation on scans of a normal subject and of an AD subject. This step was carried out and described in previous work [21,22, Appendix I]. Landmark differences across protocols that contribute significantly to volume estimates concerned the definition of the most posterior slice, the superior border, and the separation of hippocampal tissue from the parahippocampal gyrus at the level of the subiculum, along the hippocampal body [21].

A second step was designed to operationalize the numerous and fuzzy anatomic landmark differences into a limited number of well-defined units, that could undergo quantitative examination. This second step of the project is described specifically as (*i*) the operationalization of protocol differences (i.e., the definition of a limited number of three-dimensional [3D] units able to account for all interprotocol differences in a way that was sufficiently well defined to undergo quantitative investigation) and (*ii*) the estimation, for each unit, of reliability values, contribution to total hippocampal volume, and informative value regarding AD-related atrophy.

During the third step of the project, the information gathered in these previous steps was given to a panel of hippocampus experts to carry out an evidence-based Delphi procedure facilitating a consensual definition [23] of a harmonized protocol. After this definition of a harmonized protocol is achieved, a small group of "master tracers" will segment a set of benchmark images accordingly.

We plan to upload these labels on an interactive web system allowing standard learning, qualification, and periodical certification of the ability of tracers to segment according to the harmonized protocol. Tracers from participating centers, who previously segmented a set of ADNI images based on their local protocols, will then qualify and resegment the same images based on the harmonized protocol, to obtain data for the validation of the harmonized protocol (see the flowchart of the Validation phase [24]). A similar procedure will allow us to validate the protocol versus neuropathological data.

2. Methods

2.1. Operationalization of differences among protocols

This study capitalizes on previous work in which we extracted landmark variability of the 12 most prevalent hippocampal segmentation protocols in the AD literature [21, Appendix I]. We operationalized protocol differences (i.e., we collapsed these differences into a limited number of units, sufficiently well defined to lend themselves to quantitative investigation). In practice, the wide landmark heterogeneity has been reduced [21] and turned into "positive" units consisting of a finite number of elementary blocks named segmentation units (SUs). These SUs actually can be segmented as contiguous labels within the boundaries of the hippocampus on coronal MR images, and with the help of simultaneous 3D visualization, they can thus be measured and tested. SUs have also been modeled as 3D digital objects.

The definition of SU landmarks (see SU Protocol, Appendix I) was based on published landmarks definitions, drawn from the previous survey of hippocampal segmentation protocols [21] according to the following criteria: 1) definitions based on internal landmarks were preferred whenever possible, because they are more invariant to image orientation; 2) if different definitions were equally clear, the most frequently used was adopted; otherwise, the clearest definition was chosen.

We segmented SUs (M. Bocchetta and R. G.) as contiguous labels on ADNI MR images to obtain quantitative information about intra- and interrater segmentation reliability, as well as to obtain informative value regarding AD-related atrophy.

2.2. Image selection

For this study, 3D T1-weighted structural MR images at 1.5T were acquired from the ADNI database [25]. The ADNI was launched in 2003 by the National Institute on

Aging, the National Institute of Biomedical Imaging and Bioengineering, the U.S. Food and Drug Administration, private pharmaceutical companies, and not-for-profit organizations, as a US$60 million, 5-year public–private partnership. The primary goal of ADNI has been to test whether serial MR imaging, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and to monitor their effectiveness, as well as to reduce time and costs of clinical trials. The principal investigator of this initiative is Michael W. Weiner, MD (VA Medical Center and University of California, San Francisco, CA, USA). ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from more than 50 sites across the Unites States and Canada. The initial goal of ADNI was to recruit 800 adults, age 55 to 90 years old, to participate in the research—approximately 200 cognitively normal elderly individuals to be monitored for 3 years, 400 people with MCI to be monitored for 3 years, and 200 people with early AD to be monitored for 2 years. For up-to-date information, please see www.adni-info.org [26].

For the computation of reliability measures, we first selected a sample of 20 subjects (reliability sample, Table 1), four per each degree of atrophy severity evaluated using the visual medial temporal atrophy scale by Scheltens [27] (Table 1).

For the quantification of SU contribution to AD-related volume differences, a power analysis was carried out on this sample of 20 subjects to define the sample size allowing reliable computation ($\alpha = 0.05$). The required sample size was n = 77, composed of 31 control subjects, 23 MCI subjects, and 23 AD patients, which we selected from ADNI, matching for gender and age, to form the volume sample (Table 1). MCI patients were chosen among those who subsequently converted to AD, with abnormal cerebrospinal fluid (CSF) amyloid beta $(A\beta)_{1–42}$ levels ($A\beta_{1–42} < 192$ pg/mL), similar to the selected AD patients. Instead, control subjects had normal CSF $A\beta_{1–42}$ levels ($A\beta_{1–42} \geq 192$ pg/mL). CSF $A\beta_{1–42}$ values were measured from lumbar puncture using a standardized protocol, as described previously in detail [28,29].

2.3. Image processing

A combination of several freely available tools was used to prepare the ADNI MR images for manual segmentation. Medical Image Network Common Data Form images were downloaded directly from the ADNI data set [30]. Prior to tracing, the 3D images were aligned through a six-parameter registration (translations, rotations) using the Montreal Neurological Institute package AutoReg (version 0.98v) (McConnell Brain Imaging Centre, Montreal, Quebec, Canada) along the slope determined by the line that passes through the anterior commissure (AC) and the posterior commissure (PC) of the brain (AC-PC line). We used the Montreal Neurological Institute template ICBM152 Nonlinear Symmetric with 1 × 1 × 1-mm voxel dimensions as the reference. Further normalization to the template (e.g., scaling, shear) was not carried out, nor did images undergo other corrections (e.g., denoising, inhomogeneity bias field estimation), to preserve all the original signal, especially relative to the white matter (WM) layer of the alveus/fimbria. Therefore, all MR images were traced in the space defined by the AC-PC line with minimal signal change. To keep into account individual differences in brain size, the volumes

Table 1
Sociodemographic and clinical features of the and ADNI samples used for reliability (n = 20) and volume (n = 77) analyses

	Mean (SD) values			*P* value		AD + MCI	*P* value
	Control (n = 8)	MCI (n = 9)	AD (n = 3)	MCI vs control	AD vs control	(n = 12)	AD + MCI vs control
Reliability sample (n = 20)							
Age, years	76 (4)	77 (6)	78 (7)	.287	.356	77 (6)	.229
Gender, female	2 (25%)	2 (22%)	2 (67%)	1.000	.491	4 (33%)	1.000
Education, years	16 (2)	17 (2)	15 (2)	.412	.477	16 (2)	.678

	Mean (SD) values			*P* value			
	Control (n = 31)	MCI (n = 23)	AD (n = 23)	MCI vs control	AD vs control	MCI vs AD	AD + MCI vs control
Volume sample (n = 77)							
Age, years	76 (5)	76 (6)	76 (6)	.816	.704	.895	.718
Gender, female	15 (48%)	11 (48%)	11 (48%)	.999	.999	.999	.961
Education, years	16 (3)	16 (3)	15 (3)	.732	.791	.697	.511

Abbreviations: MCI, mild cognitive impairment; AD, Alzheimer's disease.

NOTE. The reliability sample included four subjects for each of the 5 degrees of atrophy on Scheltens' scale [27] and was used to assess intraclass correlation coefficients for each segmentation unit. The volume sample included control subjects with normal amyloid beta (Aβ) levels, MCI subjects who subsequently converted to AD and with abnormal Aβ levels, and AD patients with abnormal Aβ CSF levels. This sample was used for the analyses of segmentation unit volumes (Table 3).

we obtained were then corrected using the total intracranial volume (TIV) measure reported in the ADNI data set, with the formula

$$\frac{\text{Volume}}{\text{TIV}\times\text{Mean TIV}},$$

where Mean TIV is the average of the TIVs of the whole sample of 77 subjects.

2.4. Segmentation

Hippocampal SU volumes were obtained from tracings based on the protocol for the SUs (Appendix I) defined in the operationalization phase.

Each SU within the left hippocampus of the 20 ADNI images (reliability sample) was traced manually on contiguous coronal brain sections, with simultaneous visualization of the axial and sagittal planes, by two tracers (R. G. and M. Bocchetta) to compute intraclass correlation coefficients (ICCs) for each SU volume. One tracer (R. G.) was already an expert in hippocampal segmentation and was involved in the project on the harmonization of hippocampal protocols since its beginning. Her former ICC values, computed previously on an independent sample of 20 subjects taken from a local data set [31], was 0.94 for intrarater reliability and 0.89 for interrater reliability. The second tracer (M. Bocchetta) was not involved previously in the project on the harmonization of hippocampal protocols, nor did she have additional experience in hippocampal segmentation besides the learning phase, and therefore was considered as a "naive" tracer. She completed the learning of hippocampal segmentation on the same local sample of 20 images used for the learning of R. G. The ICCs for the second tracer were 0.89 for intrarater reliability and 0.88 for interrater reliability. The segmentation protocol followed by both tracers in the computation of these former ICC values was that of Pruessner and colleagues [32].

Subsequently, one rater (M. Bocchetta) traced each SU manually, according to the protocol provided in Appendix I, within the right and left hippocampi on the 77 ADNI images. The tracer was blind to diagnosis.

Tracings were performed using the interactive MultiTracer software developed at the Laboratory of Neuro Imaging at the University of California at Los Angeles (Los Angeles, CA, USA). MultiTracer allows simultaneous 3D navigation in the axial and sagittal planes, use of thin brush and subvoxel computation, and simultaneous tracings of different labels. The zoom, and hence interpolation, was kept constant (coronal view, ×5, sagittal view, ×3, axial view, ×1) and the direction of tracing was from rostral to caudal.

Hippocampi with all SUs were traced on approximately 30 contiguous, 1-mm thick coronal brain sections.

2.5. Volume computation and statistical analysis

We computed SU volumes stereologically by multiplying segmented area by slice thickness and summing these partial volumes. This was carried out by MultiTracer using the method Frust volume. This method gives better results in whole-volume interpolation when the drawn areas are not aligned completely. The volume is therefore calculated by assuming that the structure extends from the center of the first plane on which it was drawn to the center of the last plane on which it was drawn, with the square root of areas varying linearly when moving from the center of one plane to the center of the next.

We used PS version 3.0.14 (Vanderbilt University, Nashville, TN) to perform a post hoc power analysis on results from the reliability sample, considering $\alpha = .05$. The effect size was calculated as follows:

$$\text{Cohen's d}=\frac{\text{Mean of control subjects}-\text{Mean of patients}}{\text{Pooled SD}},$$

where SD is standard deviation and

$$\text{Pooled SD}=\sqrt{\frac{(n_{\text{control}}-1)\times SD^2_{\text{control}}+(n_{\text{patient}}-1)\times SD^2_{\text{patient}}}{n_{\text{control}}+n_{\text{patient}}}}.$$

We carried out *t* tests for group comparisons of SU values and ICCs with SPSS 12.0. We computed ICCs for both intra- and interrater reliability with a two-way mixed-effects model (single measure). ICC values relate to volume estimates only and not to spatial overlap (e.g., Dice or Kappa coefficient).

2.6. Modeling and rendering

We created 3D volumes using MATLAB (The Mathworks, Natick, MA, USA). First we generated a regular finely gridded volume that would enclose all subunits. We then determined the planes of this volume that matched the contour slices, then tagged the corresponding plane voxels located inside the contours with a different tag for each subunit. The resulting volume was thus, temporarily, a stack of filled contours with extra space between them. We filled these spaces using distance field interpolation [33], producing the final volume models.

To produce the renderings presented in Figs. 1 and 2 [32,34–44], we generated surfaces of the volume models using the Visualization Toolkit (http://www.vtk.org) contour filter. With Blender 3D (http://www.blender.

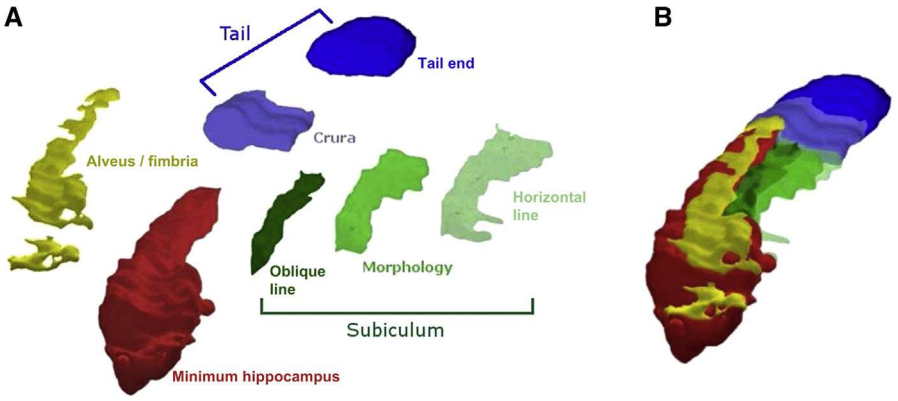

Fig. 1. Segmentation units (SUs) summarizing the variability among the 12 most common protocols for manual hippocampal segmentation in Alzheimer's disease [21]. (A) SUs are represented individually: the minimum hippocampus, alveus/fimbria, tail (divided further into two nonoverlapping subunits), and subiculum (further divided into three overlapping subunits). (B) Three-dimensional rendering of the hippocampus, resulting from fitting all SUs together. None of the 12 protocols corresponded to this shape.

org), we then smoothed the voxel-indented surfaces, applied colors and transparency, and added proper lighting for final rendering.

3. Results

During the first step of this work, we operationalized the many landmark differences among protocols (i.e., we summarized these differences into a limited number of positive portions of hippocampal tissue; Fig. 1), included or excluded differentially by the examined protocols (Fig. 2) that can be segmented independently on MR imaging and that can therefore undergo quantitative investigation (Video).

3.1. Operationalization of differences among protocols

The extraction of differences among the surveyed protocols led to the definition of four SUs: the minimum hippocampus, the alveus/fimbria, the subiculum, and the tail. Two of these SUs (the subiculum and the tail) were divided further into subunits (see SU protocol, Appendix I and video).

3.1.1. Minimum hippocampus

The minimum hippocampus corresponds to the portion of the hippocampus that is segmented by all the protocols surveyed in the previous work [21]. This includes the entire hippocampal head, separated by the amygdala through the

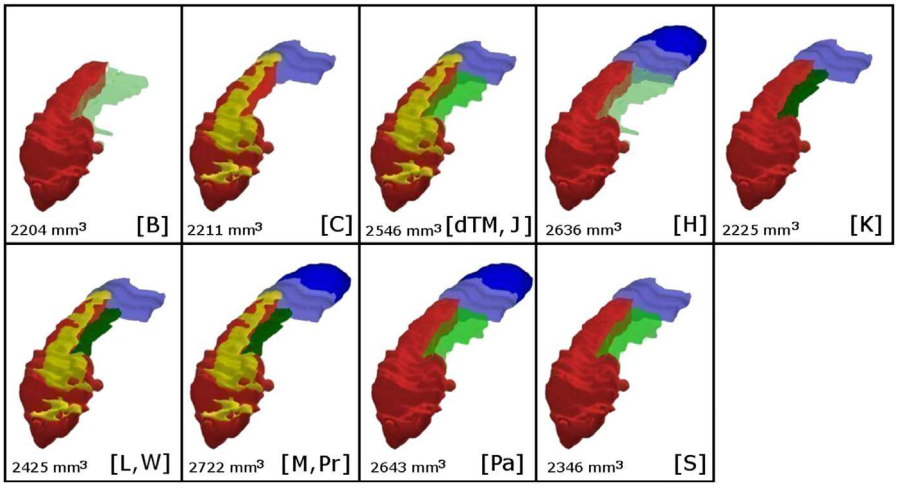

Fig. 2. Segmentation unit (SU)-based three-dimensional reconstructions of hippocampi as defined by each of the 12 surveyed segmentation protocols. Models corresponds to protocols as follows: [B], Bartzokis and colleagues [35]; [C], Convit and colleagues [34]; [dTM, J], deToledo-Morrell and colleagues [42] and Jack [41]; [H], Haller and colleagues [40]; [K], Killiany and colleagues [36]; [L, W], Lehéricy and colleagues [37] and Watson and colleagues [39]; [M, Pr], Malykhin and colleagues [38] and Pruessner and colleagues [32]; [Pa], Pantel and colleagues [43]; [S], Soininen and colleagues [44]. Numbers denote hippocampal volumes obtained by summing up the volume of those SUs included in each of the surveyed protocols [10]. SU volumes were computed on a single Alzheimer's Disease Neuroimaging Initiation (ADNI) control subject.

help of the 3D navigation tools, and part of the hippocampal body. The medial and caudal boundaries of the body are defined by the most restrictive criteria for these boundaries found among protocols. The medial boundary consists of a vertical line drawn from the most medial aspect of the cornu ammonis 1 down to the parahippocampal WM [34]. The most caudal slice was that where both superior and inferior colliculi can be visualized jointly [35]. The superior border is defined by the lower boundary of the alveus/fimbria (Fig. 1; see also the protocol of SUs, Appendix I).

3.1.2. Alveus/fimbria

This unit has been defined as the hyperintense tissue located on the dorsolateral aspect of the hippocampal head and body. The most caudal slice of alveus/fimbria is that where the superior and inferior colliculi are visualized jointly, according to the most restrictive definition of hippocampal body [35], in AC–PC-oriented images. This is the last slice where the WM can be seen as a white line parallel to the lower boundary of the hippocampal body. In subsequent slices, it diverges from the hippocampal body, bending vertically into the fornices, which are not segmented by any of the examined protocols.

3.1.3. Subiculum

At the level of the hippocampal body, the subiculum is defined by the tissue located medial to the minimum hippocampus. The different ways to segment this portion have been operationalized into three segmentation subunits: oblique line, horizontal line, and morphology. The subunit oblique line collapsed the two criteria of a 45° oblique line [32] and of an oblique line following the inclination of the parahippocampal WM [36–39] (Oblique line, Fig. 1A). The medial boundary of the oblique line consists of a line having the same slope as the parahippocampal WM below it (Fig. 1A; the protocol of SUs, Appendix I).

The medial boundary of the horizontal line is defined by a horizontal line connecting the highest point of the parahippocampal WM medially to the ambient cistern [35,40].

Last, the morphology subunit is defined by a line following the visible contour of the medial border of the hippocampus, and relying on the morphological details [41–44].

This segmentation unit was named subiculum because of its approximate location in the region of the anatomic subiculum, but does not overlap exactly with it, and it is only aimed to represent the differences in the segmentation of the medial border of the hippocampal body.

3.1.4. Tail

The tail unit corresponds to the most caudal portion of the hippocampus, adjacent to the most restrictive boundary for the minimum hippocampus [35]. The tail is divided into two subunits. Specifically, the crura subunit collapsed the two criteria prescribing to use as most caudal slice 1)the slice where the crus [34,36,37,39,42] and 2) the slice where both crura [41,44] of the fornices can be seen in their full profile. This unit is thus defined as the tissue included between the most caudal slice of the minimum hippocampus and the first slice where, in AC–PC-oriented images, and in the rostrocaudal direction, both crura of the fornices can be visualized in full profile (Fig. 1A). The second subunit, tail end, is defined as all the most caudal gray matter (GM) tissue (Fig. 1A), with exclusion of the vestigial GM as allowed by the visible differences in GM intensity and by the 3D navigation (see the protocol for the SUs, Appendix I).

As represented in Fig. 1B, the sum of all the earlier defined SUs covers the whole hippocampal tissue, according to the most inclusive definition that can be provided based on published landmarks. For a 3D animated version of this model, see Video . We defined this sum of all SUs as maximum hippocampus. In the maximum hippocampus, the tail is composed by both tail subunits (crura and tail end). In volume computations, the subiculum of the maximum hippocampus consists of the morphology subiculum subunit only, which was the largest of the three subunits, as detected by the quantitative investigation described in the following paragraph.

None of the surveyed protocols corresponded to such an inclusive definition. Instead, assembling the SUs based on landmarks definition provided in the surveyed protocols allowed the visualization of the 3D shape of the hippocampus, segmented based on each of them (Fig. 2). These 3D models show clearly how relevant landmark differences are for the differences of volume estimates across protocols.

3.2. Quantitative investigation of SUs

The next step was aimed to quantify the intra- and interrater reliability in the segmentation of each SU and of subunits, in the quantification of the contribution of each SU to the total volume of the hippocampus, and in the quantification of the information that each SU conveys about the atrophy resulting from AD.

In the reliability sample ($n = 20$), the difference between atrophy severity, as measured by the Scheltens's scale, in patients (3.0 ± 1.0) and control subjects (0.5 ± 0.5) was significant ($P < .0005$).

Reliability computations using volume-based ICC showed that all SUs can be segmented reliably, with the alveus/fimbria having the lowest ICC values (Table 2). Noticeably, the intrarater reliability of the minimum hippocampus and the alveus/fimbria traced together (intrarater reliability, 0.99; 95% confidence interval, 0.983–0.997) was significantly greater than that of the alveus/fimbria alone (intrarater reliability, 0.86; 95% confidence interval 0.687–0.944; $P < .05$). Contrary to expectations, the subiculum traced based on the morphology criterion had higher intra- and interrater reliability than the subiculum segmented using arbitrary lines (Table 2). Although the crura subunit of the tail had greater reliability figures than tail end, tail end was

Table 2
Reliability analysis

Segmentation Unit	Intrarater	Interrater
Minimum hippocampus	0.992 (0.980–0.997)	0.974 (0.936–0.990)
Alveus/fimbria	0.863 (0.687–0.944)	0.885 (0.734–0.953)
Minimum hippocampus + alveus/fimbria	0.993 (0.983–0.997)	0.973 (0.934–0.989)
Subiculum		
Oblique line	0.964 (0.911–0.985)	0.907 (0.781–0.962)
Morphology	0.981 (0.952–0.992)	0.937 (0.848–0.975)
Horizontal line	0.980 (0.951–0.992)	0.932 (0.836–0.972)
Tail		
Crura	0.998 (0.994–0.999)	0.937 (0.847–0.974)
Tail end	0.988 (0.971–0.995)	0.905 (0.775–0.961)

NOTE. Values are intraclass correlation coefficients obtained from the reliability sample (Table 1). Numbers are point estimates, with 95% confidence intervals in parentheses.

also characterized by very high ICCs (Table 2). This is noticeable, also because the vestigial tissue was excluded from this SU based on morphological information.

In the volume sample (n = 77), a significant difference was found between the CSF $A\beta_{1-42}$ levels of AD (136.0 ± 26.3) and MCI (133.7 ± 23.4) patients vs control subjects (242.7 ± 25.2; $P < .0005$ for both comparisons), but not between AD and MCI ($P = .755$).

The analysis of the contribution of SUs to total hippocampal volume and of their informative value regarding AD-related atrophy showed that the minimum hippocampus and the tail end had the highest values (Table 3). Among the subiculum subunits, the morphology subunit displayed the largest contribution and informative value regarding AD-related atrophy (Table 3). In any case, most subiculum SUs were not significantly different between patients and control subjects (with the exception of morphology for the left hippocampus, Table 3) because of their small size and thus the small effect size (0.38 for morphology, 0.33 for oblique line, and 0.36 for horizontal line). SUs not primarily affected by AD pathology, such as the alveus/fimbria, displayed up to 19% of tissue loss (Table 3).

Overall, the volume of the maximum hippocampus displayed up to 20% tissue difference in MCI, and up to 27% tissue difference in AD compared with control subjects. The minimum hippocampus and the tail were the SUs that best reflected this amount of tissue loss in patients (Table 3).

Table 3
Segmentation unit volumes and their contribution to AD-related atrophy

Segmentation Unit	Control (n = 31)	Percent of total hippocampus*	MCI (n = 23)	AD (n = 23)	Percent difference MCI vs control†	Percent difference AD vs control‡
Left hippocampus						
MinH	1467 (204)	60	1122 (263)	1023 (251)	23.5§	30§
Alveus/fimbria	248 (45)	10	232 (61)	200 (48)	6.5	19§
Subiculum	243 (72)	10	220 (84)	213 (64)	9.5	12
Oblique line	196 (67)	8	178 (66)	176 (53)	9	10
Morphology	243 (72)	10	220 (84)	213 (64)	9.5	12
Horizontal line	234 (72)	9	210 (78)	211 (62)	10	10
Tail	485 (131)	20	383 (99)	353 (101)	21¶	27§
Crura	190 (74)	8	177 (70)	146 (69)	6.5	23¶
Tail end	296 (120)	12	206 (76)	206 (86)	30¶	30¶
MaxH	2443 (291)	100	1957 (348)	1788 (342)	20§	27§
Right hippocampus						
MinH	1462 (232)	60	1214 (247)	1061 (241)	17§	27§
Alveus/fimbria	255 (47)	11	258 (71)	225 (65)	−1	12¶
Subiculum	225 (79)	9	208 (89)	184 (56)	8	18¶
Oblique line	181 (67)	8	167 (71)	150 (46)	8	17
Morphology	225 (79)	9	208 (89)	184 (56)	8	18¶
Horizontal line	220 (78)	9	203 (83)	182 (54)	7.5	17
Tail	487 (151)	20	349 (115)	349 (131)	28.5§	28.5§
Crura	187 (75)	8	169 (68)	140 (69)	10	25¶
Tail end	301 (120)	12	181 (113)	209 (110)	40§	31¶
MaxH	2429 (303)	100	2029 (372)	1820 (369)	16.5§	25§

Abbreviations: MCI, mild cognitive impairment; AD, Alzheimer's disease; MinH, minimum hippocampus (red segmentation unit in Fig. 1); MaxH, maximum hippocampus (corresponding to the most inclusive landmarks, and using the morphology criterion as the medial border of the hippocampal body; Fig. 1B).

NOTE. Numbers denote mean volume measured in cubic millimeters, corrected by total intracranial volume, and standard deviation in parentheses in the analysis sample (see Table 1).

*Proportion of segmentation units to total hippocampal volume.

†Volume differences of subjects with MCI vs control subjects.

‡Volume differences of AD patients vs control subjects.

§Significant difference at $P < .001$.

¶Significant difference at $P < .05$, *t test*.

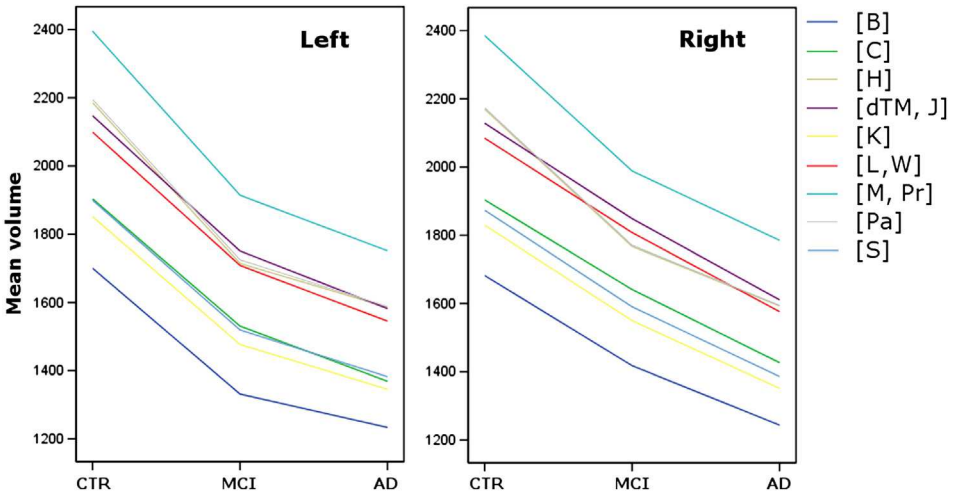

Fig. 3. Different volume estimates obtained by the 12 surveyed protocols in 31 control subjects, 23 patients with Alzheimer's disease (AD), and 23 subjects with mild cognitive impairment (MCI). Volume estimates are obtained by summing up the volumes of the segmentation units (as presented in Table 3) compounding each protocol, according to the models presented in Fig. 2. CTR, control; [B], Bartzokis and colleagues [29]; [C], Convit and colleagues [28]; [H], Haller and colleagues [34]; [dTM, J], deToledo-Morrell and colleagues [36] and Jack [35]; [K], Killiany and colleagues [30]; [L, W], Lehéricy and colleagues [31] and Watson and colleagues [33]; [M, Pr], Malykhin and colleagues [32] and Pruessner and colleagues [26]; [Pa], Pantel and colleagues [37]; [S], Soininen and colleagues [38].

3.3. SUs as sources of variability across protocols

As an empirical countercheck, the volume of the 12 surveyed protocols was obtained by summing the volumes of the SUs included in each (Fig. 2). We obtained the mean volumes of 31 control subjects, and 23 MCI and 23 AD patients for the 12 protocols. The graph shown in Fig. 3 denotes that they indeed provide different hippocampal volume estimates. Volumes ranges (mean among right and left hippocampi) are 1691 to 2390 mm^3 for control subjects, 1375 to 1951 mm^3 for MCI patients, and 1239 to 1768mm^3 for AD patients, with differences being up to 1.4-fold the volume of the most restrictive protocol across the 12 protocols that we examined.

4. Discussion

With this study, we acquired the knowledge required to carry out evidence-based harmonization of hippocampal segmentation protocols. First, we operationalized differences of landmarks among the most common AD-related hippocampal segmentation protocols into SUs. Second, we estimated the properties of SUs that will be useful to a Delphi panel to develop a harmonized protocol including all or a subset of the SUs: segmentation reliability, contribution to total hippocampal volume, and contribution to AD-related atrophy.

The operationalization phase showed that (*i*) a relatively small portion of the hippocampus (the minimum hippocampus) amounting to only 60% of the hippocampal tissue as defined by the most inclusive published landmarks is included in all the surveyed protocols and (*ii*) the wide variability among protocols can be summarized into three basic points of heterogeneity: inclusion of the alveus/fimbria, segmentation of the medial boundary of the hippocampal body, and choice of the most caudal slice for segmentation. As illustrated in Fig. 2, SUs can be assembled variably to obtain different definitions of the hippocampus, with obviously heterogeneous volumes, of which the 12 examined in a previous work [21] are a subset.

The second part of this study consisted of collecting quantitative information on SUs: reliability measures and volumetric contributions of SUs to total hippocampal volume and to AD-related atrophy. Results showed a rather homogeneous pattern of reliability values for all SUs, with the exception of the alveus/fimbria. Reliability values were rather high, which may suggest that the particularly detailed definition of landmarks that we used in this work, made necessary by the need to identify the SUs, may prompt tracers to pay attention to many more details than usual, thus leading to higher ICC values. This result clearly benefits the harmonization effort.

Estimating reliability figures of SUs was one of the main objectives of this study. Indeed, some segmentation protocols exclude some portions of the hippocampus on the grounds that they would introduce exceedingly high error. For example, the most caudal slices of the hippocampus, excluded by a number of protocols [34–37,39,41,42,44] is often believed to lead to higher variability because it may be difficult to separate hippocampal and vestigial tissue. Similarly, arbitrary lines are often preferred to the morphology criterion when segmenting the medial border of the hippocampal body, because these are believed to achieve more stable results. The results of this study falsified the latter belief showing that the use of the morphology criterion for the segmentation of the medial

boundary of the hippocampal body is associated with high intra- and interrater reliability figures. The exclusion of the WM of the alveus/fimbria may be a more risky option because it leads to the lowest reliability values, whereas the segmentation of the entire hippocampal body, including the alveus/fimbria, achieves a much greater reliability (intrarater reliability, 0.993; interrater reliability, 0.973; Table 2). We also showed that small SUs, such as the subiculum, or SUs that cannot be considered as hippocampus proper, such as the alveus/fimbria, do convey information about AD-related atrophy, and thus may add sensitivity to a harmonized protocol for use in AD.

Delphi panelists who will express their opinion on inclusion or exclusion of SUs in a harmonized protocol will be able to weigh the quantitative information of this study in the light of anatomic–physiological knowledge. For example, they may wish to weigh the increased segmentation reliability associated with inclusion of alveus/fimbria, and the increased sensitivity of the protocol to AD-related atrophy vs the fact that the alveus/fimbria is not the hippocampus proper.

The entire effort, aimed to achieve a consensual harmonized protocol, will provide the gold standard for the manual segmentation of the hippocampus for use in AD studies and clinical applications. Algorithms for automated segmentation are, of course, an elective tool for a less time-consuming hippocampal segmentation. A further advantage of these algorithms consists in their high reliability, although different algorithms are also based on different landmark definitions, leading to incomparable results. The harmonized protocol may also be a standard for automated segmentation algorithms to which to conform to segment the hippocampus reliably for scientific and clinical purposes in the field of AD.

This work has shown that a harmonization effort can lead plausibly to concrete results with a reasonable amount of work. Of course, there are important limitations. We collected quantitative data on a sample of 77 ADNI subjects. This group size was computed based on the preliminary results on a pilot 20 subjects sample (see Methods). Nonetheless, a much larger sample would be necessary to detect significant differences among patient groups for the smallest SUs, such as the three subiculum subunits. Moreover, measurements of the tail end subunit were taken excluding the vestigial tissue based on the description provided by the atlases by Mai and colleagues [45] and Duvernoy [46]. Nonetheless, we have not created or measured an SU corresponding to the vestigial tissue, which may facilitate the achievement of a consensus on the inclusion of the most caudal slices of the hippocampus. Last, the ICC values were based on volumes only, and may underestimate differences in spatial overlap. They were also computed based on the segmentations of two raters who belong to the same laboratory. Later in this project, segmentations of SUs will be available from more tracers and from different laboratories, and will allow more stable ICC computations. The certification platform will also take spatial overlap into consideration for new rater approval.

The next step of this project will consist of the use of this information in the context of recursive rounds of the Delphi panel; panelists will access quantitative information, 3D models, literature, and the anonymous answers of other panelists to weigh pros and cons of the inclusion of each SU, and achieve an evidence-based consensus on landmark definition for a harmonized protocol for hippocampal segmentation.

Acknowledgments

The Alzheimer's Association has provided logistic support for the update meetings of the project in Toronto (April 2010), Honolulu (July 2010 and April 2011), and Paris (July 2011). Wyeth, part of the Pfizer group, and Lilly have provided unrestricted grants in support of the work reported in this article. A follow-up project has been funded by the Alzheimer's Association: A Harmonized Protocol for Hippocampal Volumetry: An EADC-ADNI Effort (grant no. IIRG -10-174022). The project principle investigator (PI) is Giovanni B. Frisoni, IRCCS Fatebenefratelli, Brescia, Italy; the co-PI is Clifford R. Jack, Mayo Clinic, Rochester, MN; the Statistical Working Group is led by Simon Duchesne, Laval University, Quebec City, Canada; and project coordinator is Marina Boccardi, IRCCS Fatebenefratelli, Brescia, Italy. European Alzheimer's Disease Consortium (EADC) centers (local principal investigators) are IRCCS Fatebenefratelli, Brescia, Italy (G. B. Frisoni); University of Kuopio and Kuopio University Hospital, Kuopio, Finland (H. Soininen); Höpital Salpètriere, Paris, France (B. Dubois and S. Lehericy); University of Frankfurt, Frankfurt, Germany (H. Hampel); University Rostock, Rostock, Germany (S. Teipel); Karolinska Institutet, Stockholm, Sweden (L.-O. Wahlund); Department of Psychiatry Research, Zurich, Switzerland (C. Hock); Alzheimer Centre, Vrije University Medical Centre, Amsterdam, The Netherlands (F. Barkhof and P. Scheltens); the Dementia Research Group Institute of Neurology, London, UK (N. Fox); and NEUROMED, Centre for Neuroimaging Sciences, London, UK (A. Simmons). Alzheimer's Disease Neuroimaging Initiative (ADNI) centers are Mayo Clinic, Rochester, MN (C. R. Jack); University of California at Davis, Davis, CA (C. DeCarli); University of California at Los Angeles (UCLA), Los Angeles, CA (G. Bartzokis); University of California at San Francisco (UCSF), San Francisco, CA (M. Weiner and S. Mueller); Laboratory of NeuroImaging, UCLA, Los Angeles, CA (P. M. Thompson); Rush University Medical Center, Chicago, IL (L. deToledo-Morrell); Rush Alzheimer's Disease Center, Chicago, IL (D. Bennet); Northwestern University, Chicago, IL (J. Csernansky); Boston University School of Medicine, Boston, MA (R. Killiany); John Hopkins University, Baltimore, MD (M. Albert); the Center for Brain Health, New York, NY (M. De Leon); and Oregon Health & Science University, Portland, OR (J. Kaye). Other centers are McGill University, Montreal, Quebec, Canada (J.

Pruessner); University of Alberta, Edmonton, Alberta, Canada (R. Camicioli and N. Malykhin); Department of Psychiatry, Psychosomatic, Medicine & Psychotherapy, Johann, Wolfgang Goethe-University, Frankfurt, Germany (J. Pantel); Wayne State University, Detroit, MI (C. Watson); and the Institute for Ageing and Health, Wolfson Research Centre, Newcastle General Hospital, Newcastle, UK (J. O'Brien). Population-based studies include PATH Through Life, Australia (P. Sachdev and J. J. Maller); the SMART-Medea Study, The Netherlands (M. I. Geerlings); and the Rotterdam Scan Study, The Netherlands (T. denHeijer). The Statistical Working Group includes the Fatebenefratelli Association for Biomedical Research, San Giovanni Calibita - Fatebenefratelli Hospital, Rome, Italy (P. Pasqualetti); Laval University, Quebec City, Canada (S. Duchesne); and the Montreal Neurological Institute, McGill University, Montreal, Canada (L. Collins). The advisor for clinical issues include P. J. Visser, Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands; EADC PIs include B. Winbald, Karolinska Institute, Sweden; and Ll Froelich, Central Institute of Mental Health, Mannheim, Germany. Advisors for dissemination and education include G. Waldemar, Copenhagen University Hospital, Copenhagen, Denmark; the ADNI PI is M. Weiner, UCSF, San Francisco, CA. Advisors for population Studies include L. Launer, National Institute on Aging (NIA), Bethesda, MD; and W. Jagust, University of California, Berkeley, CA.

The magnetic resonance images used in this paper belong to the ADNI data set. Data collection and sharing for this project was funded by the ADNI (National Institutes of Health grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. 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 Rev March 26, 2012 coordinated by the Alzheimer's Disease Cooperative Study at the University of California at San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles.

This research was also supported by National Institutes of Health grants P30 AG010129 and K01 AG030514.

N. R. and S. D. have received funding support from the Ministère du Développement Économique, de l'Innovation et de l'Exportation du Québec.

We thank Leyla deToledo-Morrell for the language revision.

References

[1] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:263–9.

[2] Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol 2007; 6:734–46.

[3] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:280–92.

[4] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:270–9.

[5] Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, et al. Revising the definition of Alzheimer's disease: a new lexicon. Lancet Neurol 2010;9:1118–27.

[6] Collins DL, Pruessner JC. Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. Neuroimage 2010; 52:1355–66.

[7] Rodionov R, Chupin M, Williams E, Hammers A, Kesavadas C, Lemieux L. Evaluation of atlas-based segmentation of hippocampi in healthy humans. Magn Reson Imaging 2009;27:1104–9.

[8] Kennedy KM, Erickson KI, Rodrigue KM, Voss MW, Colcombe SJ, Kramer AF, et al. Age-related differences in regional brain volumes: a comparison of optimized voxel-based morphometry to manual volumetry. Neurobiol Aging 2009;30:1657–76.

[9] Hasan KM, Pedraza O. Improving the reliability of manual and automated methods for hippocampal and amygdala volume measurements. Neuroimage 2009;48:497–8.

[10] Bosscher L, Scheltens P. Functional imaging in dementia. In: Qizilbash N, ed. Evidence-based dementia practice. Oxford, UK: Blackwell; 2002. p. 162–9.

[11] Duchesne S, Pruessner J, Collins DL. Appearance-based segmentation of medial temporal lobe structures. Neuroimage 2002;17:515–31.

[12] Barnes J, Foster J, Boyes RG, Pepple T, Moore EK, Schott JM, et al. A comparison of methods for the automated calculation of volumes and atrophy rates in the hippocampus. Neuroimage 2008; 40:1655–71.

[13] Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, et al. Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease, mild cognitive impairment, and elderly control subjects. Neuroimage 2008;43:59–68.

[14] Brewer JB, Magda S, Airriess C, Smith ME. Fully-automated quantification of regional brain volumes for improved detection of focal atrophy in Alzheimer disease. AJNR Am J Neuroradiol 2009; 30:578–80.

[15] Colliot O, Chetelat G, Chupin M, Desgranges B, Magnin B, Benali H, et al. Discrimination between Alzheimer disease, mild cognitive

impairment, and normal aging by using automated segmentation of the hippocampus. Radiology 2008;248:194–201.
[16] Geuze E, Vermetten E, Bremner JD. MR-based in vivo hippocampal volumetrics: 1. Review of methodologies currently employed. Mol Psychiatry 2005;10:147–59.
[17] Konrad C, Ukas T, Nebel C, Arolt V, Toga AW, Narr KL. Defining the human hippocampus in cerebral magnetic resonance images: an overview of current segmentation protocols. Neuroimage 2009; 47:1185–95.
[18] Jack CR Jr, Barkhof F, Bernstein MA, Cantillon M, Cole PE, Decarli C, et al. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. Alzheimers Dement 2011; 7:474–85.
[19] Frisoni GB, Jack CR. Harmonization of magnetic resonance-based manual hippocampal segmentation: a mandatory step for wide clinical use. Alzheimers Dement 2011;7:171–4.
[20] A Harmonized Protocol For Hippocampal Volumetry: An EADC-ADNI Effort. Available at: http://www.hippocampal-protocol.net/SOPs/index.html. Accessed April 9, 2013.
[21] Boccardi M, Ganzola R, Bocchetta M, Pievani M, Redolfi A, Bartzokis G, et al. Survey of protocols for the manual segmentation of the hippocampus: preparatory steps towards a joint EADC-ADNI harmonized protocol. J Alzheimer Dis 2011;26(Suppl 3)61–75.
[22] A Harmonized Protocol For Hippocampal Volumetry: An EADC-ADNI Effort. Surveyed Protocols. Available at: http://www.hippocampal-protocol.net/SOPs/investigatedprotocols.html. Accessed April 9, 2013.
[23] Murphy M, Black N, Lamping D, McKee C, Sanderson C, Askham J, et al. Consensus development methods, and their use in clinical guideline development. Health Technol Assess 1998;2:1–88.
[24] A Harmonized Protocol For Hippocampal Volumetry: An EADC-ADNI Effort. Flow Chart Validation. Available at http://www.hippocampal-protocol.net/SOPs/LINK_PAGE/Flow_chart_validation_phase.pdf. Accessed April 9, 2013.
[25] Alzheimer's Disease Neuroimaging Initiative. Available at: http://www.adni-info.org. Accessed April 9, 2013.
[26] Alzheimer's Disease Neuroimaging Initiative. Sharing Alzheimer's Research Data with the World. Available at: http://adni.loni.ucla.edu. Accessed April 9, 2013.
[27] Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P, et al. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. J Neurol Neurosurg Psychiatry 1992; 55:967–72.
[28] Olsson A, Vanderstichele H, Andreasen N, De Meyer G, Wallin A, Holmberg B, et al. Simultaneous measurement of beta-amyloid(1–42), total tau, and phosphorylated tau (Thr181) in cerebrospinal fluid by the xMAP technology. Clin Chem 2005;51:336–45.
[29] Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, et al. Cerebrospinal fluid biomarker signature in Alzheimer's Disease Neuroimaging Initiative subjects. Ann Neurol 2009;65:403–13.
[30] Alzheimer's Disease Neuroimaging Initiative (ADNI). Data and Samples. Available at: http://adni.loni.usc.edu/data-samples. Accessed March 3, 2014.
[31] Galluzzi S, Testa C, Boccardi M, Bresciani L, Benussi L, Ghidoni R, et al. The Italian Brain Normative Archive of structural MR scans: norms for medial temporal atrophy and white matter lesions. Aging Clin Exp Res 2009;21:266–76.
[32] Pruessner JC, Li LM, Serles W, Pruessner M, Collins DL, Kabani N, et al. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. Cereb Cortex 2000;10:433–42.
[33] Marker J, Braude I, Museth K, Breen D. Contour-based surface reconstruction using implicit curve fitting, and distance field filtering and interpolation. Proc Int Workshop Volume Graphics 2006;95–102.
[34] Convit A, De Leon MJ, Tarshish C, De Santi S, Tsui W, Rusinek H, et al. Specific hippocampal volume reductions in individuals at risk for Alzheimer's disease. Neurobiol Aging 1997;18:131–8.
[35] Bartzokis G, Altshuler LL, Greider T, Curran J, Keen B, Dixon WJ. Reliability of medial temporal lobe volume measurements using reformatted 3D images. Psychiatry Res 1998;82:11–24.
[36] Killiany RJ, Moss MB, Albert MS, Sandor T, Tieman J, Jolesz F. Temporal lobe regions on magnetic resonance imaging identify patients with early Alzheimer's disease. Arch Neurol 1993; 50:949–54.
[37] Lehericy S, Baulac M, Chiras J, Pierot L, Martin N, Pillon B, et al. Amygdalohippocampal MR volume measurements in the early stages of Alzheimer disease. AJNR Am J Neuroradiol 1994; 15:929–37.
[38] Malykhin NV, Bouchard TP, Ogilvie CJ, Coupland NJ, Seres P, Camicioli R. Three-dimensional volumetric analysis and reconstruction of amygdala and hippocampal head, body and tail. Psychiatry Res 2007;155:155–65.
[39] Watson C, Jack CR Jr, Cendes F. Volumetric magnetic resonance imaging: clinical applications and contributions to the understanding of temporal lobe epilepsy. Arch Neurol 1997;54:1521–31.
[40] Haller JW, Banerjee A, Christensen GE, Gado M, Joshi S, Miller MI, et al. Three-dimensional hippocampal MR morphometry with high-dimensional transformation of a neuroanatomic atlas. Radiology 1997;202:504–10.
[41] Jack CR Jr. MRI-based hippocampal volume measurements in epilepsy. Epilepsia 1994;35:S21–9.
[42] deToledo-Morrell L, Stoub TR, Bulgakova M, Wilson RS, Bennett DA, Leurgans S, et al. MRI-derived entorhinal volume is a good predictor of conversion from MCI to AD. Neurobiol Aging 2004;25:1197–203.
[43] Pantel J, O'Leary DS, Cretsinger K, Bockholt HJ, Keefe H, Magnotta VA, et al. A new method for the in vivo volumetric measurement of the human hippocampus with high neuroanatomical accuracy. Hippocampus 2000;10:752–8.
[44] Soininen HS, Partanen K, Pitkanen A, Vainio P, Hanninen T, Hallikainen M, et al. Volumetric MRI analysis of the amygdala and the hippocampus in subjects with age-associated memory impairment: correlation to visual and verbal memory. Neurology 1994;44:1660–8.
[45] Mai JK, Assheuer J, Paxinos G. Atlas of the human brain. San Diego, CA: Academic Press; 1997.
[46] Duvernoy HM. The human hippocampus: functional anatomy, vascularization and serial sections with MRI. 3rd ed. Berlin: Springer; 1998.