NeuroImage 000 (2021) 117931

Contents lists available at ScienceDirect

NeuroImage

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

Hippocampal subfield volumes across the healthy lifespan and the effects of MR sequence on estimates

Aurélie Bussy[a,b,*], Eric Plitman[a,c], Raihaan Patel[a,d], Stephanie Tullo[a,b], Alyssa Salaciak[a], Saashi A. Bedford[a,c], Sarah Farzin[a], Marie-Lise Béland[a], Vanessa Valiquette[a,b], Christina Kazazian[a], Christine Lucas Tardif[d,e,f], Gabriel A. Devenyi[a,c], M. Mallar Chakravarty[a,b,c,d,*], Alzheimer's Disease Neuroimaging Initiative

[a] Computional Brain Anatomy (CoBrA) Laboratory, Cerebral Imaging Centre, Douglas Mental Health University Institute, Montreal, QC, Canada
[b] Integrated Program in Neuroscience, McGill University, Montreal, Canada
[c] Department of Psychiatry, McGill University, Montreal, QC, Canada
[d] Department of Biological and Biomedical Engineering, McGill University, Montreal, QC, Canada
[e] McConnell Brain Imaging Centre, Montreal Neurological Institute, Montreal, QC, Canada
[f] Department of Neurology and Neurosurgery, McGill University, Montreal, QC, Canada

A B S T R A C T

The hippocampus has been extensively studied in various neuropsychiatric disorders throughout the lifespan. However, inconsistent results have been reported with respect to which subfield volumes are most related to age. Here, we investigate whether these discrepancies may be explained by experimental design differences that exist between studies. Multiple datasets were used to collect 1690 magnetic resonance scans from healthy individuals aged 18–95 years old. Standard T1-weighted (T1w; MPRAGE sequence, 1 mm^3 voxels), high-resolution T2-weighted (T2w; SPACE sequence, 0.64 mm^3 voxels) and slab T2-weighted (Slab; 2D turbo spin echo, $0.4 \times 0.4 \times 2$ mm^3 voxels) images were included. The MAGeT Brain algorithm was used for segmentation of the hippocampal grey matter (GM) subfields and perihippocampal white matter (WM) subregions. Linear mixed-effect models and Akaike information criterion were used to examine linear, second or third order natural splines relationship between hippocampal volumes and age. We demonstrated that stratum radiatum/lacunosum/moleculare and fornix subregions expressed the highest relative volumetric decrease, while the cornus ammonis 1 presented a relative volumetric preservation of its volume with age. We also found that volumes extracted from slab images demonstrated different age-related relationships compared to volumes extracted from T1w and T2w images. The current work suggests that although T1w, T2w and slab derived subfield volumetric outputs are largely homologous, modality choice plays a meaningful role in the volumetric estimation of the hippocampal subfields.

1. Introduction

Medial temporal lobe structures, particularly the hippocampus, have been extensively studied for their involvement in various neuropsychiatric disorders such as Alzheimer's disease (Bobinski et al., 2000; Braak and Braak, 1991; Pol et al., 2006; Zhao et al., 2019), schizophrenia (Dwork, 1997; Heckers, 2001; Nelson et al., 1997), major depression disorder (Campbell and MacQueen, 2004; Malykhin et al., 2010; Stockmeier et al., 2004), and frontotemporal dementia (Laakso et al., 2000; Muñoz-Ruiz et al., 2012). To better contextualize group differences in case-control studies, an understanding of normative variation of hippocampal structure across the adult lifespan is critical. This information is crucial in order to better understand how deviations from this trajectory may precede the frank onset (or even the prodromes) of various neuropsychiatric disorders. Unfortunately, current studies investigating the relationship between hippocampus structure and age have rendered inconsistent results.

An examination of studies that report upon the relationship between hippocampal volume and age in healthy aging highlights these inconsistencies: some studies report global preservation (Good et al., 2001; Sullivan et al., 2005, 1995), while others report overall reduction (Bussy et al., 2019; Kurth et al., 2017; Malykhin et al., 2017; Raz et al., 2004). More recently, several studies have begun to characterize this relationship at the level of the grey matter (GM) hippocampal subfields, including cornu ammonis (CA)1, 2, 3 and 4, subiculum and stratum radiatum/lacunosum/moleculare (SRLM), using innovative new techniques (Goubran et al., 2013; Olsen et al., 2019; Pipitone et al., 2014; Winterburn et al., 2013; Yushkevich et al., 2015). In most studies reviewed in de Flores et al. (de Flores et al., 2015), the CA1 and subiculum appear to be most impacted by aging. However, in other recent studies, the CA1 (Amaral et al., 2018; Voineskos et al., 2015) and subiculum (Daugherty et al., 2016) have been shown to be relatively preserved. In the present work, we suggest that the observed discrepancies in findings may be explained by the study design dispar-

* Corresponding authors at: Computional Brain Anatomy (CoBrA) Laboratory, Cerebral Imaging Centre, Douglas Mental Health University Institute, Montreal, QC, Canada.
E-mail addresses: aureliee.bussy@gmail.com (A. Bussy), mallar.chak@gmail.com (M.M. Chakravarty).

https://doi.org/10.1016/j.neuroimage.2021.117931
Available online xxx

ities - across a range of methodological choices - that exist between studies.

First, a key methodological difference is the anatomical definition of the hippocampal subfields. This issue is controversial and thus, different atlases are used in the literature (Amaral et al., 2018; Iglesias et al., 2015; Palombo et al., 2013; Winterburn et al., 2015; Yushkevich et al., 2015; Yushkevich et al., 2015). In order to facilitate the segmentation of small subfields, protocols use diverse definitions of subfields, which include CA1 combined with CA2 (Bender et al., 2018), CA3 combined with dentate gyrus (DG) (de Flores et al., 2015), CA3 combined with CA4 and DG (Shing et al., 2011) or CA4 solely combined with DG (Winterburn et al., 2015; Wisse et al., 2012). Second, various segmentation protocols are employed to study the subfields, including manual delineation (La Joie et al., 2010; Mueller et al., 2007), semi-automated approaches (Yushkevich et al., 2010), and automated methods (Iglesias et al., 2015; Pipitone et al., 2014; Van Leemput et al., 2009; Yushkevich et al., 2015).

Of note, a recent effort has been started by the Hippocampal Subfields Group (http://www.hippocampalsubfields.com/) to develop a harmonized protocol based on expert consensus and histological evidence as a means to facilitate comparison of findings across subject groups (Olsen et al., 2019; Wisse et al., 2017). However, despite these significant and important efforts, there remains two additional aspects related to study design that are outside the scope of segmentation protocols, namely: 1) the non-linear relationship of the hippocampal subfield volumes with age, 2) the magnetic resonance image (MRI) acquisition protocols. Further, to completely characterize the age-related trends of the hippocampal circuitry, it is critical to examine the examination of peri-hippocampal white matter (WM) subregions such as alveus, fornix, fimbria and mammillary bodies (MB). The current manuscript primarily focuses on the impact of these two components.

Importantly, previous studies have examined diverse participant age ranges. While some cohorts consider the entire adult lifespan (Amaral et al., 2018; de Flores et al., 2015; Mueller et al., 2007; Zheng et al., 2018), others only assess subjects older than 65 (Frisoni et al., 2008; Wisse et al., 2014). These variations can lead to inconsistencies in findings, since it is well-accepted that the relationship between brain structures and age is typically non-linear (Coupé et al., 2017; Tullo et al., 2019); thus, sampling a smaller age range can lead to conclusions that should not be used to interpret data outside that range. Nonetheless, only a few studies have investigated the non-linear relationships between age and hippocampal subfield volumes (de Flores et al., 2015; Malykhin et al., 2017; Mueller et al., 2007; Ziegler et al., 2012).

MRI protocols are highly heterogeneous in the literature and its impact on hippocampal subfield definitions is still not clear. Indeed, studies have used standard T1-weighted (T1w) images (Amaral et al., 2018; Chételat et al., 2008; Frisoni et al., 2008), more specialized T2-weighted (T2w) images (Mueller et al., 2007; Wisse et al., 2014) or proton-density-weighted (PDw) images (de Flores, La Joie, and Chételat, 2015; La Joie et al., 2010 R.). Further, the resolution used to examine hippocampal subfields is highly variable across these image acquisitions. For example, studies used isotropic whole brain scans at 0.7 mm^3, 1 mm^3 (Pereira et al., 2014; Wisse et al., 2014) (typically for T1w images, and rarely T2w) or $0.78 \times 0.78 \times 1.5$ mm^3 resolution (Voineskos et al., 2015), or "slab" scans at $0.4 \times 0.4 \times 2$ mm^3 resolution (Mueller and Weiner, 2009) (typically for T2w and PD images). The latter is acquired in a region of interest that covers the amygdala and hippocampus either partially or entirely from its anterior to posterior extent. Taken together, these different acquisition parameters may contribute to volume estimation differences potentially explaining the variation observed in age-related subfield relationships. Further, there are very few studies that used high-resolution isotropic scans which allow for the visualization of the molecular layers (i.e. "dark band") of the hippocampus, often considered a prerequisite landmark for manual or automated identification of hippocampal subfields (Eriksson et al., 2008; Goubran et al., 2014; Winterburn et al., 2013; Wisse et al., 2017). Also, there have been suggestions in several studies that slab T2w or PDw scans should be preferentially used in the study of hippocampal subfields to obtain coronal high resolution (La Joie et al., 2010; Yushkevich et al., 2015). Concurrently, there has also been minimal investigation with respect to the use of higher resolution isotropic images for this purpose (Wisse et al., 2014, 2012). Finally, given that the hippocampus and its subfields make up a critical circuit in the brain's memory network, some groups (Iglesias et al., 2015), including ours (Amaral et al., 2018; Tardif et al., 2018), have begun to include the peri-hippocampal WM in studies examining the hippocampal subfields. This is a critical step forward towards examining this specific anatomy at the circuit level.

Given the inconsistency in the literature, there is a clear need to better characterize age-related relationships between the volume of the hippocampal subfields and peri-hippocampal WM subregions while performing a critical assessment of experimental design choices. Here, we examine the role of image acquisition (including standard isotropic T1w, high-resolution isotropic T2w and slab hippocampus-specific T2w) on hippocampal subfield volumes, while assessing the reliability of the hippocampal subfield measures.

2. Methods

2.1. Image acquisition and participants

2.1.1. Image acquisition types

Here, we examine three different scan types from multiple sources. These data come from multiple datasets collected by our group and publicly available datasets (described further in Section 2.1.2). These scan types were targeted because two of them (standard T1w MPRAGE and slab T2-weighted 2D turbo spin echo [TSE]) have been commonly used in the hippocampal subfield literature. The final one, a high-resolution whole brain T2w acquisition, is introduced here as a methodology that potentially addresses some of the limitations with respect to resolution and field-of-view.

- *T1-weighted:* This is the protocol most commonly used in neuroimaging studies and has been extensively employed to study the hippocampus, among other structures (Pereira et al., 2014). Here, we used the Alzheimer's Disease Neuroimaging Initiative (ADNI) magnetization prepared - rapid gradient echo (MPRAGE) protocol (Jack et al., 2008; Mugler and Brookeman, 1990), since these parameters have been commonly used to investigate hippocampus physiology and pathology.
- *Slab T2-weighted 2D TSE:* This protocol uses the same ADNI oblique acquisition with 2 mm thick slices perpendicular to the long axis of the hippocampus and 0.4 mm x 0.4 mm in the coronal plane. Recently, this sequence became the standard procedure to study the hippocampus subfields. Users typically take advantage of the high-resolution of the coronal plane to segment the subfields while sacrificing resolution through the anterior-posterior direction. While the type of scanner field has been investigated in regard to the hippocampal measurements (Derix et al., 2014; Marques and Norris, 2018; Parekh et al., 2015), it is still unclear what the design trade-off of the anisotropic scans does in terms of consistency and precision of the measurement.
- *High-resolution 0.64 mm^3 T2-weighted:* This protocol uses the SPACE sequence, a 3D TSE sequence with slab selective, variable excitation pulse. It has been developed in our laboratory for the purpose of increasing the resolution and subfield contrast in the hippocampus. Furthermore, this sequence provides isotropic images across the whole brain and allows for the use of standard linear and nonlinear registration methods that expect whole-brain coverage (Chakravarty et al., 2013; Klein et al., 2009).

Table 1
Demographics by dataset: Final sample after motion and segmentation quality control (QC) across *Healthy Aging (HA), Alzheimer's Disease Biomarkers (ADB), Alzheimer's Disease Neuroimaging Initiative (ADNI), Cambridge Centre for Ageing and Neuroscience (Cam-CAN)* and *Test-retest datatsets.* ***np*** *= number of participants;* ***ns*** *= number of scans.*

	ADB (np=52) (ns=100)	*ADNI (np=135) (ns=233)*	*Cam-CAN (np=376) (ns=376)*	*HA (np=84) (ns=167)*	*Test-retest (np=18) (ns=54)*	*Overall (np=665) (ns=930)*
Age						
Mean (SD)	69.3 (5.7)	71.6 (6.9)	49.9 (17.2)	46.4 (16.1)	26.8 (5.6)	55.4 (18.5)
Median [Min, Max]	69.5 [56.0, 81.0]	70.0 [56.0, 93.0]	49.0 [18.0, 85.0]	46.0 [18.0, 80.0]	26.0 [20.0, 42.0]	61.0 [18.0, 93.0]
Sex						
F	61 (61.0%)	160 (68.7%)	196 (52.1%)	93 (55.7%)	33 (61.1%)	543 (58.4%)
M	39 (39.0%)	73 (31.3%)	180 (47.9%)	74 (44.3%)	21 (38.9%)	387 (41.6%)
Sequence						
T1	52 (52.0%)	135 (57.9%)	376 (100.0%)	84 (50.3%)	18 (33.0%)	665 (71.5%)
T2	48 (48.0%)	0 (0.0%)	0 (0.0%)	83 (49.7%)	18 (33.0%)	149 (16.0%)
Slab	*0 (0.0%)*	*98 (42.1%)*	*0 (0.0%)*	*0 (0.0%)*	*18 (33.0%)*	*116 (12.5%)*

2.1.2. Datasets

The five datasets used to examine the relationship of the hippocampal subfields with age are outlined below, and all include some variation of the acquisition methods described in the previous section. All the datasets collected by our group at the Douglas Mental Health University Institute were approved by its Research Ethics Board.

- *Healthy Aging (HA)*. This dataset was collected and scanned on a 3T Siemens Trio MRI scanner using a 32-channel head coil at the Douglas Mental Health University Institute, Montreal, Quebec, Canada, and contains 111 participants aged 18–80 (Tullo et al., 2019). We analyzed two types of MRI images: standard MPRAGE (1 mm^3) and high-resolution T2-weighted (TSE; 0.64 mm^3).
- *Alzheimer's Disease Biomarkers (ADB)*. This dataset was also collected and scanned on a 3T Siemens Trio MRI scanner using a 32-channel head coil at the Douglas Mental Health University Institute, Montreal, Quebec, Canada, (Tullo et al., 2019). From this study, we used 68 healthy elderly participants (56–81). The same acquisition protocol as the HA dataset was used, providing standard MPRAGE (1 mm^3) and high-resolution T2-weighted (TSE sequence; 0.64 mm^3).
- *Alzheimer's Disease Neuroimaging Initiative (ADNI)*. ADNI is a publicly available multicenter study from which we included 317 healthy participants, aged 56–95, scanned on 3T General Electric, Philips or Siemens scanners, depending on the acquisition's site. We used two types of MRI images: standard MPRAGE (1 mm^3) (Jack et al., 2008) and slab T2-weighted 2D TSE with 2 mm thick slices perpendicular to the long axis of the hippocampus and 0.4 mm dimensions in the coronal plane (Mueller et al., 2007, 2018; Thomas et al., 2004).
- *Cambridge Centre for Ageing and Neuroscience (Cam-CAN)*. Cam-CAN is a large-scale dataset collecting MRI scans at the Medical Research Council Cognition and Brain Sciences Unit in Cambridge, England, using a 3T Siemens TIM Trio scanner with a 32-channel head coil (Shafto et al., 2014; Taylor et al., 2017). We included 652 healthy individuals (aged 18–88) with standard MPRAGE (1 mm^3).
- *Test-retest*. Test-retest dataset was collected at the Douglas Mental Health University Institute, Montreal, Canada. Twenty-one healthy participants aged 20–42 were recruited and scanned using a 3T Siemens PRISMA scanner. These participants underwent three different MR sequence acquisitions that reproduce the three different types of acquisitions described above: standard MPRAGE (1 mm^3), high-resolution T2-weighted (TSE sequence; 0.64 mm^3) and slab T2-weighted 2D TSE sequence ($0.4 \times 0.4 \times 2$ mm^3).

The total number of scans after quality control (QC) of the images (see *Section 3 Image processing*), was 930, and consisted of individuals aged 18–93 (58.4% female) (Table 1; Demographics of the initially included participants before scan QC and preprocessing summarized in supplementary Table 1). Complete demographics and QC inclusion/exclusion criteria are available in Supplementary material; age distribution of the participants included in the analyses can be found in Supplementary figure 1. The main acquisition parameters are provided in Table 2 and more details given in Supplementary material - Acquisition parameters.

2.2. Image processing

2.2.1. Raw quality control

Structural MR images are particularly sensitive to subject motion, often resulting from involuntary movements (e.g. cardiac or respiratory motion, and drift over time). The effects of motion, including blurring and ringing, negatively impact the quality of structural MRI data (Bellon et al., 1986; Reuter et al., 2015; Smith and Nayak, 2010). Quality control (QC) of all raw images was performed by a rater (AB) using the QC procedure previously developed in our laboratory ((Bedford et al., 2020); https://github.com/CoBrALab/documentation/wiki/Motion-Quality-Control-Manual).

2.2.2. Preprocessing

The minc-bpipe-library pipeline (https://github.com/CobraLab/minc-bpipe-library) was employed to preprocess and standardize T1w images using N4 bias field correction (Tustison et al., 2010), registration to MNI space using bestlinreg (Collins et al., 1994; Dadar et al., 2018), standardization of the field-of-view and orientation of the brain using an inverse-affine transformation of a MNI space head mask, and extraction of the brain using BEaST (Eskildsen et al., 2012). The pipeline produces brain masks and quality control images to quickly evaluate all steps of the pipeline.

T2w and slab images preprocessing consists of the following steps: rigid registration of T1w to T2w or slab scan (Collins et al., 1994; Dadar et al., 2018), application of the transform file to the T1w brain masks obtained by the minc-bpipe-library pipeline to create T2w or slab brain masks, N4 correction, and extraction of T2w or slab brains using T2w or slab brain masks.

2.2.3. Automated hippocampus segmentation

The Multiple Automatically Generated Templates (MAGeT) Brain algorithm, a modified automated multi-atlas technique (Chakravarty et al., 2013; Pipitone et al., 2014), was used for segmentation of the hippocampal subfields. The MAGeT brain algorithm and documentation related to its use are available on https://github.com/CobraLab/documentation/wiki. This method uses a set of five high-quality atlases, manually segmented on 0.3 mm isotropic T1w and T2w brains, as input. We used the Winterburn et al. (2013) definitions of the GM subfields (CA1, combined CA2 and CA3 [CA2CA3], combined CA4 and dentate gyrus [CA4DG], SRLM and subiculum) and the Amaral et al. (2018) definitions of WM subregions

Table 2
Scanning parameters of the different sequences in the different datasets. TE= echo time, TR= repetition time, TI = inversion time, α= flip angle. More details in Supplementary material - Acquisition parameters.

Datasets	Contrast	Sequence	TE (ms)	TR (ms)	TI (ms)	α	Matrix	Resolution (mm^3)	Scan time (mm:ss)
HA	T1w	MPRAGE	2.98	2300	900	9	256 × 240 × 176	1	05:12
ADB	T1w	MPRAGE	2.98	2300	900	9	256 × 240 × 176	1	05:12
ADNI	T1w	MPRAGE	min full echo	2300	900	9	256 × 240 × 208	1	06:20
Cam-CAN	T1w	MPRAGE	2.99	2250	900	9	256 × 240 × 192	1	04:32
Test/retest	T1w	MPRAGE	2.01	2300	900	9	256 × 240 × 208	1	05:12
HA	T2w	SPACE	198	2500	/	/	350 × 350 × 263	0.64	13:22
ADB	T2w	SPACE	198	2500	/	/	350 × 350 × 263	0.64	10:02
Test/retest	T2w	SPACE	198	2500	/	/	350 × 350 × 263	0.64	07:35
ADNI	T2w	T2w 2D TSE	50	8020	/	150	175 × 175 × 60	0.39 × 0.39 × 2	04:20
Test/retest	T2w	T2w 2D TSE	76	8020	/	150	150 × 150 × 60	0.39 × 0.39 × 2	06:34

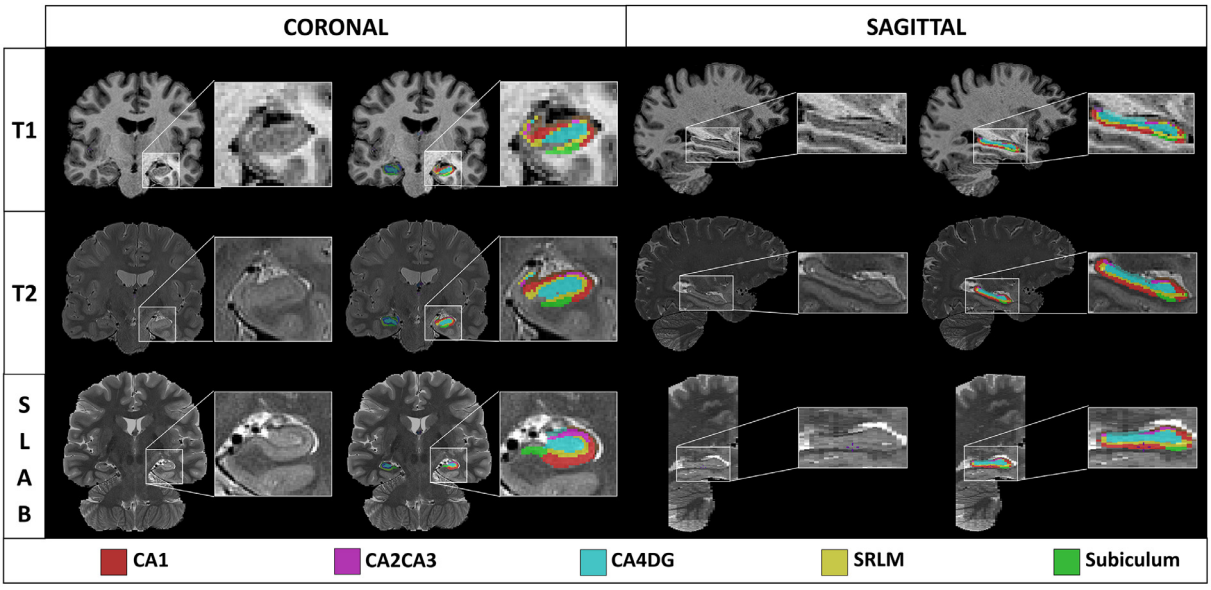

Fig. 1. Example of coronal and sagittal views of a participant's scans: T1w (1 mm^3), T2 (0.64 mm^3) and slab (0.4 × 0.4 × 2 mm) without and with the labels obtained from our segmentation protocol.

(fimbria, fornix, alveus and MB) of the hippocampus (Fig. 1). Specifically, the atlases and the corresponding T1w images have been used to segment T1w images while the atlases and the corresponding T2w images have been used to segment T2w isotropic images and T2w slab images. We used either T1w or T2w images to reduce, to the best of our ability, the bias of having one sequence type with higher contrast and intensity similarity with the image brains of the atlases.

For each sequence-type within a specific dataset (e.g. high-resolution T2w data in the ADB dataset), we first ran a "best template selection" stage (https://github.com/CoBrALab/documentation/wiki/Best-Templates-for-MAGeT) in order to select the 21 subjects with highest quality atlas-to-template segmentation. This best template selection is a MAGeT run for which we only perform the registration between the five atlases and all the subjects. The atlases-to-subjects segmentations obtained from this run are carefully quality controlled to select the 21 subjects who give the best initial segmentation. Then, the full MAGeT run is completed using these 21 subjects as a template library to segment all the subjects (see supplementary figure 3 for a schematic illustration of the full MAGeT-Brain segmentation run). Furthermore, to the best of our abilities, the 21 templates were balanced in age and sex in the same way that the entire dataset of interest, to obtain a representative set of templates. This step computationally increases the number of atlases to 105 candidate labels (21 templates x 5 atlases) specific to each dataset (since derived from the subjects of interest). Finally, the 105 candidate labels for each subject were fused using a majority vote technique to obtain final labels, in order to improve segmentation by reducing error propagation compared to traditional atlas-based segmentation procedures (Chakravarty et al., 2013; Iosifescu et al., 1997; Makowski et al., 2018; Pipitone et al., 2014; Svarer et al., 2005). MAGeT Brain uses affine and SyN nonlinear registration, which are registration options provided as part of the Advanced Normalization Tools [ANTS; (Avants et al., 2008)]. Segmentations were conducted using a focused region-of-interest (ROI) based registration step for each hemisphere independently. ROI masks were generated by converting all labels to a single value mask and then dilated with a 3 mm radius kernel in order to focus both affine and nonlinear registration, a method which reduces computational time and improves segmentation accuracy (Chakravarty et al., 2008, 2009).

For the slab segmentation, in both the "best template selection" or full MAGeT runs, the initial bulk alignment of the slab images were facilitated using their corresponding whole brain T1w images (since the small field-of-view is not optimal for accurate registrations). First, within-subject affine registrations were performed between slab and whole brain T1w images by using the label masks to handle mismatched field-of-view. Then, affine registrations were performed using the whole brain T1w images between atlases/templates and templates/subjects. Finally, nonlinear registrations were computed between the slabs images to obtain precise hippocampi segmentation.

The rater (AB) quality controlled the final labels by visual inspection following the QC procedure implemented by our group (https://github.com/CobraLab/documentation/wiki/MAGeT-Brain-

Quality-Control-(QC)-Guide) to only include high quality segmentation in the statistical analyses.

2.3. Statistical analysis

2.3.1. Relationship between age and normalized hippocampal subfield volumes

To draw general conclusions on the association between each subfield volume with age, we first combined all the datasets together in order to account for the variability in image resolution, sequence, and age range commonly encountered in the literature (Yushkevich et al., 2015). Linear mixed-effect models (lmer from *lmerTest_3.1–0* package in R 3.6.1) and natural spline (ns from *splines* package) were used to examine the relationship between hippocampal volumes and age. To study different age relationships for each structure of interest, the Akaike information criterion [AIC; (Akaike, 1974)] was used to find the relative quality of each statistical model using either linear, second, or third order natural splines. To minimize the loss of information, the model with the lowest AIC was considered the best fit for the data (Mazerolle, 2006).

Sex was used as a fixed effect, and dataset (ADB, ADNI, HA, CamCAN, Test-retest), MRI sequence (T1, T2, slab), and subject were modelled as random effects for all statistical analyses. In addition, intracranial volume (ICV) was used to account for interindividual variability in head size (Daugherty et al., 2016; de Flores et al., 2015). We used z-scored ICV (*scale* function) to help the fit of the model (1).

$$Volume \sim ns(Age, n) + Sex + scale(ICV) + (1|Sequence) + (1|Dataset) + (1|Subject) \quad (1)$$

To determine the extent to which each subfield participates in global hippocampus atrophy with age, we then assessed the relationship between each GM or WM subfield volume and age, while covarying for total hippocampal GM volume (CA1 + CA2CA3 + CA4DG + Subiculum + SRLM volumes) or WM volume (fimbria + fornix + MB + alveus volumes), respectively, as well as ICV, (2) and (3).

$$VolumeGM \sim ns(Age, n) + Sex + scale(ICV) + scale(HIPGM) + (1|Sequence) + (1|Dataset) + (1|Subject) \quad (2)$$

or

$$VolumeWM \sim ns(Age, n) + Sex + scale(ICV) + scale(HIPWM) + (1|Sequence) + (1|Dataset) + (1|Subject) \quad (3)$$

For each analysis described above, we used a Bonferroni correction to correct for multiple comparisons across our 18 structures (five GM subfields and four WM subregions per hemisphere), at a $p<0.05$ threshold for significance, resulting in a significance level of $p<0.0028$ (only corrected p-values reported throughout this paper).

To visualize these results, we used two different techniques. We first used these coefficients to create the predicted volume divided by the predicted volume at age 18 for a subject of mean ICV and mean ipsilateral hippocampal GM or WM volume when applicable. This allows us to assess the age relative volume change of each subfield with respect to its baseline volume (Figs. 2, 3, 4 and Supplementary figures 7 and 8). The second type of representation was done using the fixed effect coefficients of the statistical models described above to create the corresponding relationship of the volume with age (Supplementary figures 4, 5, 6, 8 and 10). Of note, both visualization techniques show similar shapes with the first relating these slopes in percent volume change with age whereas the second illustrates volume change with age.

Apolipoprotein E $\varepsilon 4$ allele (*APOE4*) gene is often studied in aging population since it is associated with higher risk of both early-onset and late-onset sporadic AD (Corder et al., 1993) as well as age-related cognitive impairment (Rawle et al., 2018). While some authors described that *APOE4* allele was associated with hippocampal, amygdalar and entorhinal cortex atrophy (Cherbuin et al., 2007), others did not find any structural differences between *APOE4* carriers and noncarriers (Habes et al., 2016) or even that *APOE4* noncarriers had more pronounced age-related atrophy (Bussy et al., 2019; Gonneaud et al., 2016). Although not the primary interest of this paper, age-related *APOE4* effect was investigated in a subset of 203 genotyped participants (52 ADB-T1, 48 ADB-T2, 52 HA-T1 and 51 HA-T2) with models investigating the interaction of *APOE4* with age. No *APOE4* effect was found to be related to hippocampal volume with age. Therefore, *APOE4* was discarded in our following analyses.

2.3.2. Impact of sequence type on the relationship between hippocampal subfield volumes with age

We studied the impact of the sequence type on the relationship between hippocampal subfield volumes with age *(4)*. Here, we examined whether the intercept and/or slope of the predicted model was influenced by sequence-type after covarying by sex, ICV, and ipsilateral hippocampal GM or WM volume as fixed effects, and dataset and subjects as random effects.

$$Volume \sim ns(I(Age - 18), n) * Sequence + Sex + scale(ICV) + scale(HIPGM) + (1|Dataset) + (1|Subject) \quad (4)$$

Fig. 5 and Supplementary figure 12 illustrate the relationship between age and hippocampal subfield volumes when extracting from T2w and slab images, compared to when extracting using T1w data. To visualize that, we used model coefficients to predict T1, T2w, and slab subfield volumes divided by the predicted volume at age 18 for a subject of mean ICV, and mean ipsilateral hippocampal GM or WM volume extracted from T1w images. The goal of dividing by the predicted volume at age 18 extracted from T1w images, regardless of sequence, was to remove the potential effect of over- or under-estimation of a specific sequence and thus to focus on the age-related relationship differences. Supplementary figures 11 and 13 represent the significant over- or under-estimation of the volumes, in addition to the significant age-related relationships encountered when using solely model coefficients to predict T1, T2w, and slab subfield volumes.

2.3.3. Impact of sequence type on volume estimates

Given the use of various MRI parameters to study the hippocampus subfields in the literature, we used the Test-retest dataset to compare the volume estimates from T1w, T2w, and slab sequences in the same participants to see how these diverse parameters impact subfield volumes estimation. A dependent 2-group Wilcoxon signed rank test was performed to compare the volume estimates from T1w, T2w, and slab sequences. Here, we used a Bonferroni correction significance level adjusted for 54 multiple comparisons (18 subregions x 3 sequence types), resulting in a significance level of $p < 0.00093$. Intraclass correlation coefficients (ICCs; *psych_1.8.12* package) were determined to reflect the degree of consistency (ICC [3.1]) between the volume estimates of the different sequences. ICCs were interpreted according to previously established criteria: “excellent”: 1.00–0.75, “good”: 0.74–0.60, “fair”: 0.59–0.40, and "poor": 0.39–0.00 (Cicchetti et al. 1994). The mean percentage volume difference was determined to calculate the extent of the differences in the volume estimations for each sequence.

2.3.4. Contrast to noise ratio (CNR)

Effect of age on the hippocampal subfields contrast has been demonstrated in a study (Knight et al., 2016). Therefore, we developed a protocol to assess the CNR in our data. We defined small masks in the CA1 and in the CA4DG of the hippocampal head, and in the temporal lobe white matter (Winterburn et al., 2013). Supplementary figure 15 illustrates how each mask was defined as ten consecutive voxels.

First, we used T1w and high-resolution T2w scans on HA dataset to investigate the CNR relationship with age between 20 and 80. We randomly chose four participants (two females and two males) per decade. Second, we used T1w and T2w slab scans on ADNI dataset to investigate the CNR relationship with age between 60 and 90. Here, we randomly chose four participants (two females and two males) per 5 year bin.

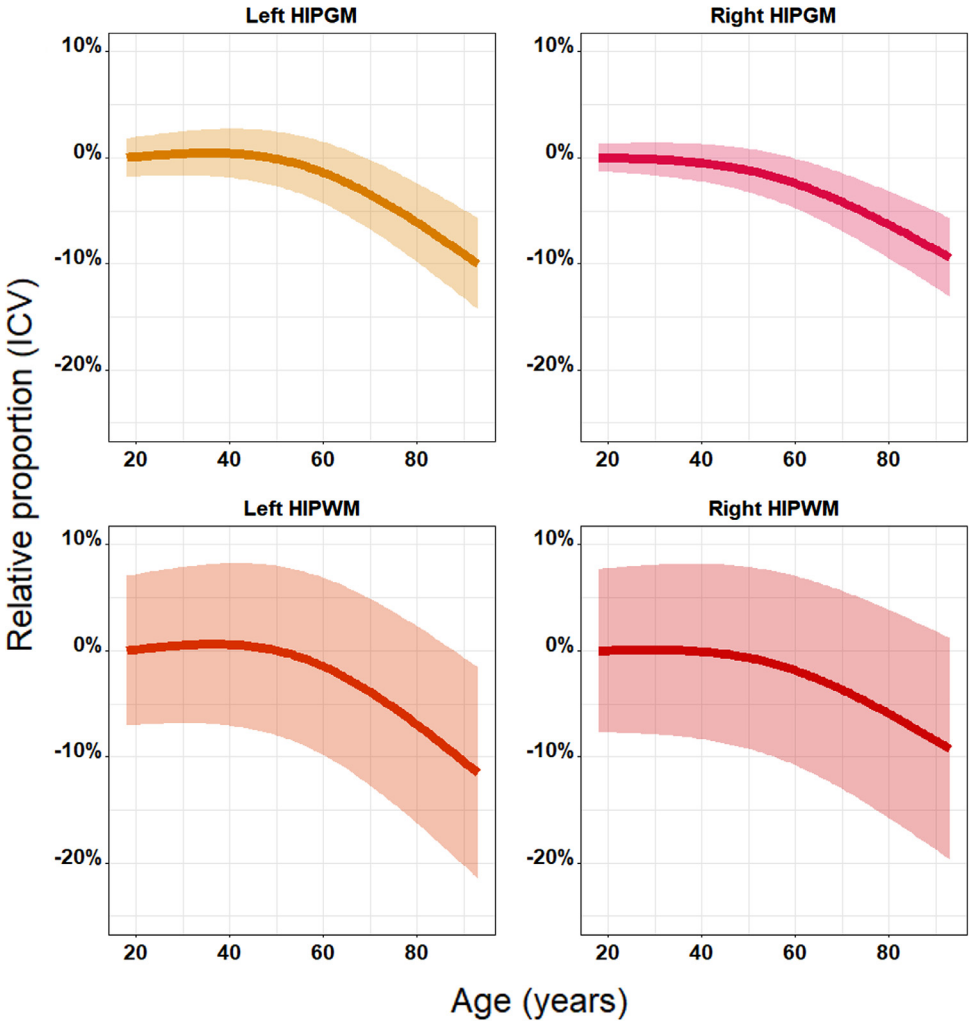

Fig. 2. Best fit models showing the relationships between age and the relative proportion of the hippocampus using the predicted volume at age 18 for a subject of mean ICV as baseline (same model with volume instead of relative proportion in Supplementary figure 3). Best fit models displayed for each structure covaried by ICV and sex as fixed effects and dataset, sequence, and subjects as random effects. Second order relationships were found to be the best fit model for all the structures: right HIPGM ($p = 3.33 \times 10^{-4}$), right HIPWM ($p = 1.05 \times 10^{-3}$), left HIPGM ($p = 2.27 \times 10^{-5}$) and left HIPWM ($p = 1.18 \times 10^{-5}$).

The CNR was calculated using the following formula (Knight et al., 2016; Winterburn et al., 2013):

$$\mathrm{CNR} = \frac{|WM - GM|}{\sqrt{\mathrm{var(GM)} + \mathrm{var(WM)}}}$$

where WM and GM represent the mean WM and GM intensities within the mask, respectively and var(WM) and var(GM) represent the variance of the two tissue classes within the mask. Linear models were used to investigate the effect of age and Bonferroni correction was used to correct for multiple comparisons.

2.3.5. HIPS comparison

HIPpocampus subfield Segmentation (HIPS; (Manjon et al., 2020; Romero et al., 2017) was applied to all the T1w scans of our datasets to investigate the differences of segmentation technique with MAGeT. We decided to investigate this other segmentation technique since it uses the same atlases as MAGeT (Winterburn et al., 2013). HIPS segmentation has been performed thanks to volBrain website (www.volbrain.upv.es), labels were manually quality controlled with our standard procedure and 689 individuals with good labels with both MAGeT and HIPS segmentations were analyzed. Linear mixed effect models investigating the interaction between the segmentation technique (MAGeT vs HIPS) and age were performed (see supplementary figure 17). Sex, ICV and right or left total hippocampal GM volume were used as a fixed effect, and dataset and subject were modelled as random effects for the statistical analyses. AIC was used to find the relative quality of each statistical model and demonstrated that all subfields were best explained using a third order natural splines. Bonferroni correction was applied on 10 structures (five bilateral GM subfields), at a $p < 0.05$ threshold for significance, resulting in a significance level of $p < 0.005$.

2.3.6. ASHS comparison

Automatic segmentation of hippocampal subfields [ASHS; (Yushkevich et al., 2015)] was applied to the 98 slab scans of the ADNI dataset to investigate the differences of volume estimation compared to MAGeT outputs. We decided to investigate this segmentation technique since ASHS segmentation protocol uses slab-specific atlases with anisotropic resolution. The atlases themselves were developed using serial histological data (Adler et al., 2014; Duvernoy, 2005; Pluta et al., 2012). The goal was to examine whether MAGeT segmentation on the slab images performed properly despite the mismatched resolution between our atlases and the slab images. Here, the absolute hippocampal volumes were expected to be different from MAGeT and ASHS, since the atlases of each technique define the hippocampal subfield boundaries differently (Yushkevich et al., 2015). However, the aim was to investigate whether the segmentation techniques gave similar age-relationships. Linear mixed effect models including the interaction between the segmentation technique (MAGeT vs ASHS) and age were performed (see supplementary figure 18). Sex, ICV and right or left total hippocampal GM volume were used as a fixed effect, and subject was modelled as random effect for the statistical analyses. AIC was used to find the relative quality of each statistical model and demonstrated that all subfields were best explained using third order natural splines. Bonferroni correction was applied on the 8 structures [bilateral CA1,

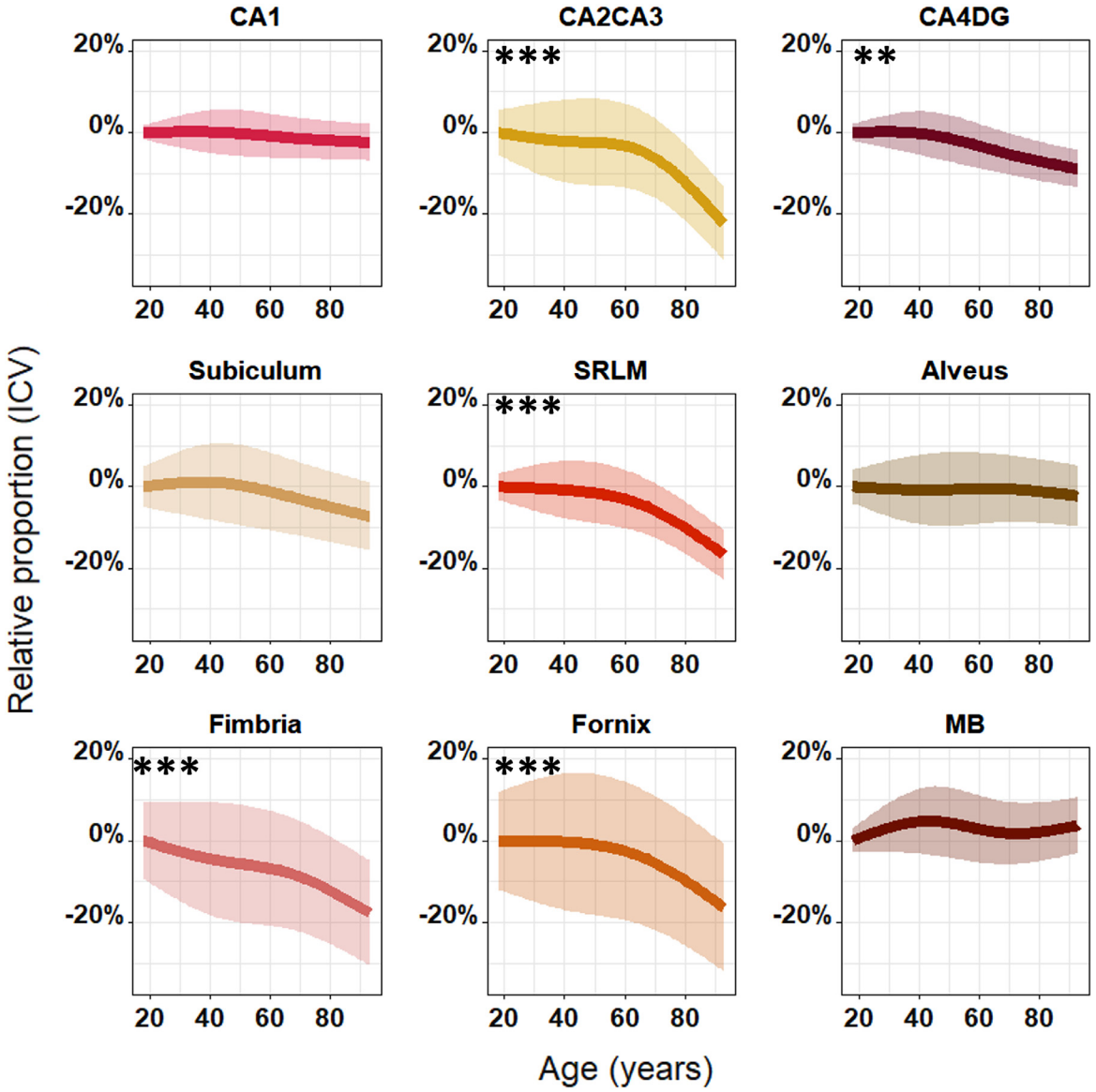

Fig. 3. Best fit models showing the relationships between age and the relative proportion of the right hippocampal subfields, using the predicted volumes at age 18 for a subject of mean ICV as baseline (same model with volume instead of relative proportion in Supplementary figure 5). Best fit models displayed for each subfield covaried by ICV and sex as fixed effects and dataset, sequence, and subject as random effects. Significant monotonic decreases were found for the right CA2CA3 ($p = 3.55 \times 10^{-7}$), CA4DG ($p = 4.28 \times 10^{-3}$), SRLM ($p = 7.74 \times 10^{-7}$), fimbria ($p = 9.45 \times 10^{-4}$) and fornix ($p = 1.79 \times 10^{-4}$). Similar relationships were found in the left hemisphere (Supplementary figures 7 and 8) * $p < 0.05$; ** $p < 0.01$ and *** $p < 0.001$ after Bonferroni correction.

CA2CA3 (defined as CA2+CA3 volumes in ASHS), CA4DG (defined as DG in ASHS) and Subiculum], at a $p < 0.05$ threshold for significance, resulting in a significance level of $p < 0.00625$.

3. Results

3.1. Relationship between hippocampal subfield volumes and age

3.1.1. All datasets normalized by the ICV

After covarying for ICV and assessing linear, second- and third-order relationships with age using the AIC, second-order models demonstrated to be the best fit for bilateral GM and WM hippocampus (Fig. 2): right HIPGM ($p = 3.33 \times 10^{-4}$), right HIPWM ($p = 1.05 \times 10^{-3}$), left HIPGM ($p = 2.27 \times 10^{-5}$) and left HIPWM ($p = 1.18 \times 10^{-5}$). These results show that, when accounting for head size difference, the relative volumes of the bilateral hippocampi were reduced by approximately 10% between age 18 and 93, with a steeper decline after age 50.

To test which subfields are involved in this 10% relative volume decrease of the hippocampi with age, individual subfield relationships with age were examined. Third-order models were observed to be the best fit for all hippocampal subfields bilaterally (Fig. 3 and Supplementary figure 5). Significant monotonic decreases were found for the right CA2CA3 ($p = 3.55 \times 10^{-7}$), CA4DG ($p = 4.28 \times 10^{-3}$), SRLM ($p = 7.74 \times 10^{-7}$), fimbria ($p = 9.45 \times 10^{-4}$) and fornix ($p = 1.79 \times 10^{-4}$). These results indicate that relative volumes of the CA2CA3, SRLM, fimbria and fornix decreased by approximately 20%, while CA4DG decreased by 10% between age 18 and 93. Interestingly, volumetric impairments seem to start earlier in CA4DG, SRLM and fimbria compared to the CA2CA3 and fornix, which seem relatively preserved until age 60. Similar relationships were found in the left hemisphere (Supplementary figures 7 and 8). In contrast, the bilateral CA1, subiculum, alveus and MB did not express significant relationships with age, and therefore were not implicated in the global volume decrease of the hippocampus.

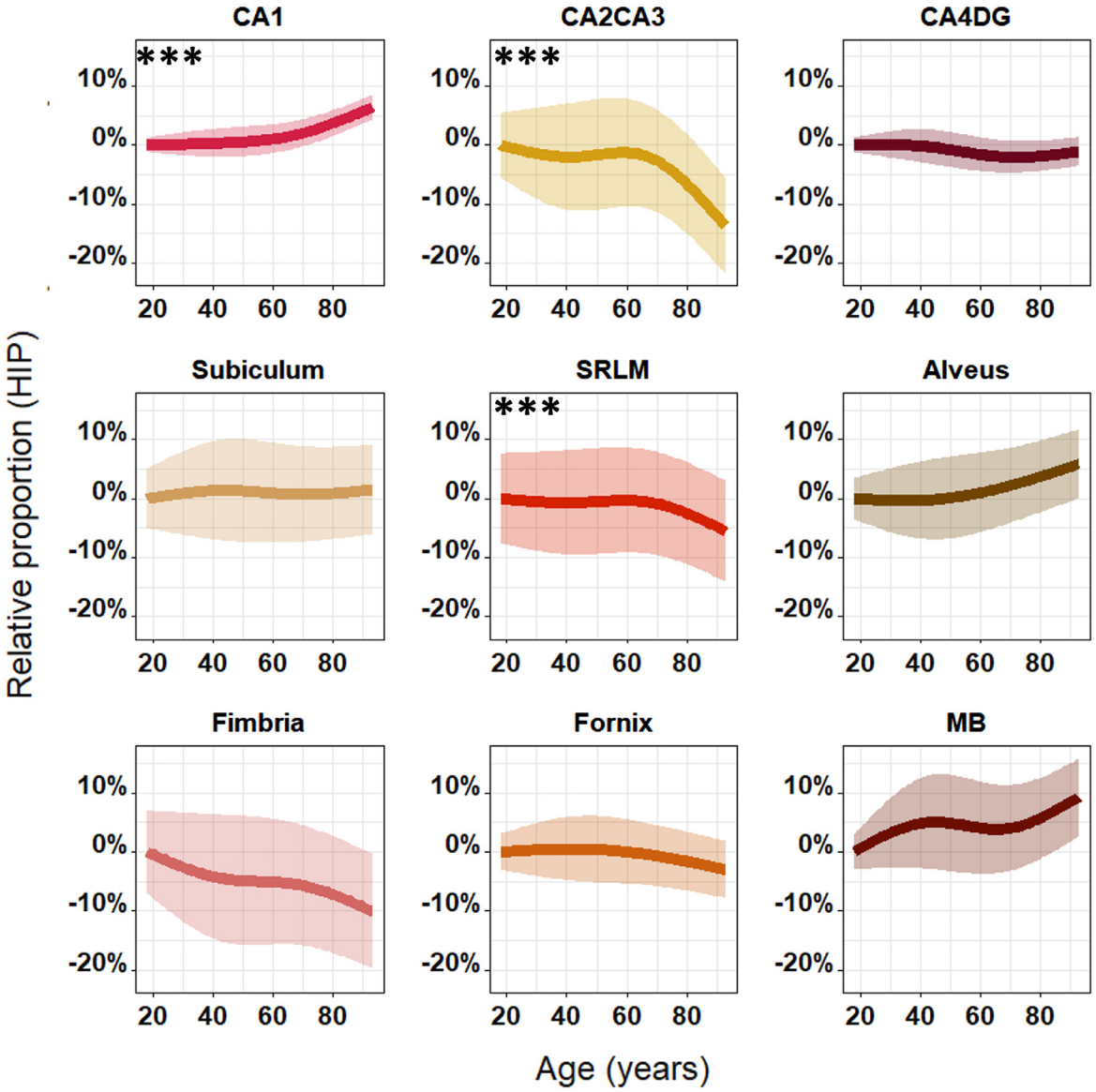

Fig. 4. Best fit models showing the relationships between age and the relative proportion of the right hippocampal subfields, using the predicted volumes at age 18 for a subject of mean ICV and mean right hippocampal GM or WM volume as baseline (same model with volume instead of relative proportion in Supplementary figure 6). Best fit model displayed for each subfield covaried by right hippocampal GM or WM volume, ICV and sex as fixed effects and dataset, sequence, and subject as random effects. Significant monotonic increases were found for the right CA1 ($p = 1.19 \times 10^{-10}$) and significant monotonic decreases were found for the right CA2CA3 ($p = 3.19 \times 10^{-4}$) and SRLM ($p = 5.29 \times 10^{-6}$). Similar investigations made in the left hemisphere can be found in Supplementary figures 9 and 10. * $p < 0.05$; ** $p < 0.01$ and *** $p < 0.001$ after Bonferroni correction.

3.1.2. All datasets normalized by the hippocampal volume

In the right hemisphere, when covarying for ipsilateral hippocampal GM or WM volume and ICV, all hippocampal subfield volumes were shown to have a best fit third-order relationship with age. Significant monotonic increases were found for the relative proportion of the right CA1 ($p = 1.19 \times 10^{-10}$). The relative volume of CA1 demonstrated a 7% increase between age 18 and 93. Significant monotonic decreases were found for the right CA2CA3 ($p = 3.19 \times 10^{-4}$) and SRLM ($p = 5.29 \times 10^{-6}$) (Fig. 4 and Supplementary figure 6). These results indicate that relative volumes of the CA2CA3 decreased by approximately 15% and SRLM by 5% between age 18 and 93. No significant volume interaction with age was found for the right CA4DG, subiculum, alveus, fimbria, fornix and MB. Similar relationships were found in the left hemisphere (Supplementary figures 9 and 10).

3.1.3. Impact of sequence on subfield volume relationship with age

T1w volumes demonstrated significant linear increases for CA1 ($p = 2.65 \times 10^{-6}$), decreases for fimbria ($p = 9.85 \times 10^{-3}$) and third-order relationships for CA2CA3 ($p = 6.59 \times 10^{-5}$) and SRLM ($p = 1.33 \times 10^{-5}$) highlighting a steeper decline after age 60 for these subfields (Fig. 5A). Compared to T1w, significant differences were found with slab for the CA1 ($p = 0.0102$) and MB ($p = 3.47 \times 10^{-14}$), demonstrating a steeper increase with age, while CA4DG ($p = 1.93 \times 10^{-3}$) and fornix ($p = 9.41 \times 10^{-7}$) manifested a stronger decline with age. Volumes estimated from the T2w sequence exhibited similar relationships with age when compared to volumes estimated from T1w, except that T2w expressed steeper increase for CA1 ($p = 5.49 \times 10^{-4}$). In Fig. 5A, the different age relationships for each sequence are represented without intercept difference to emphasize the effect of age on the estimation. While not plotted, we see in Table 5B (and in more details in Supplementary figure 11) that slab intercepts were smaller for CA1, CA2CA3, subicu-

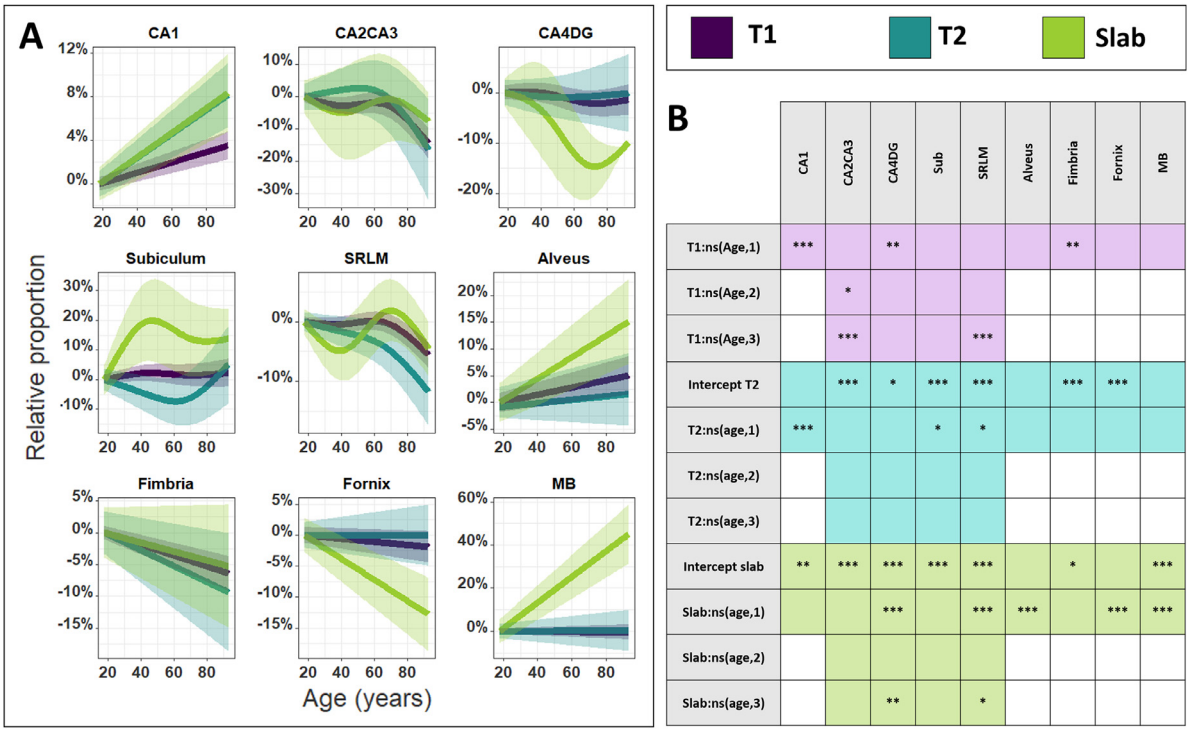

Fig. 5. A. Estimated fixed effect plots showing the relationships, separated by sequence-type, between age and the relative proportion of the right hippocampal subfields, using the predicted volume at age 18 for a subject of mean ICV and mean ipsilateral hippocampal GM or WM volume extracted from T1w images as baseline. Best fit model displayed for each subfield covaried by ipsilateral hippocampal GM or WM volume, ICV, and sex as fixed effects and dataset, sequence, and subject as random effects. **B.** Table describing the significant coefficients for the different relationships with age and the intercepts. T1w was used as reference sequence in the model and demonstrated a significant linear increase for CA1 ($p = 2.65 \times 10^{-6}$), decrease for fimbria ($p = 9.85 \times 10^{-3}$) and third order decrease for CA2CA3 ($p = 6.59 \times 10^{-5}$) and SRLM ($p = 1.33 \times 10^{-5}$). Significant differences were found with slab compared to T1w, with the CA1 ($p = 0.0102$) and MB ($p = 3.47 \times 10^{-14}$), demonstrating a steeper increase with age, while CA4DG ($p = 1.93 \times 10^{-3}$) and fornix ($p = 9.41 \times 10^{-7}$) showing a steeper decline with age. Best fit models estimated from the T2w sequence exhibited similar relationships with age than the models obtained with T1w images except that T2w expressed steeper increases with age for CA1 ($p = 5.49 \times 10^{-4}$). Supplementary figure 11 described in more details the significant intercept differences. Similar investigations made in the left hemisphere can be found in the Supplementary figures 12 and 13. * $p < 0.05$; ** $p < 0.01$ and *** $p < 0.001$ after Bonferroni correction.

lum, fimbria,and MB, while they were higher for CA4DG and SRLM. In addition, compared to T1w volumes, T2w volumes expressed smaller intercepts for CA2CA3, subiculum, and fimbria and higher intercepts for CA4DG, SRLM and fornix. Similar relationships were found in the left hemisphere (Supplementary figures 12 and 13).

3.2. Impact of different sequence on hippocampal subfield volumes estimates

We compared the volume estimates of the right hippocampal subfields obtained from T1w, T2w, and slab sequences within the same subjects (Fig. 6). Significant differences were obtained between T1w and slab sequences for all the subfields except the SRLM. These results demonstrated that T1w images lead to larger volume estimates than the slab sequence, on average by 11.2% (Supplementary Table 2), except for the CA4DG which is larger using slab images. T2w images appeared to render similar volumes compared to T1w images, except for the CA2CA3, which was larger when extracted from T1w. Moreover, T2w volumes estimates were on average 8.1% larger than the volumes extracted from slab sequences. These findings suggest that, while we cannot rule which sequence provides the most accurate volume estimation due to a lack of ground truth, slab images appear to estimate smaller volumes compared to other acquisitions

Finally, ICC (3,1) was used to calculate the consistency between each dataset volume estimates (Fig. 7). High ICC values were mostly found between T1w and slab sequences. We can also observe that the left hippocampus revealed lower ICC in most subfields. Subregions exhibiting poor consistency included left fimbria and fornix.

3.3. CNR analysis

No significant age effect was found on the CNR of the CA1 and CA4DG using the T1w, high-resolution T2w or slab T2w images of the HA or ADNI datasets (supplementary Figure 16).

3.4. HIPS segmentation analysis

Hippocampal subfield volumes extracted from HIPS or MAGeT segmentation techniques were investigated with age. Subtle significant differences in the right CA1 and in the left CA4DG were found (supplementary Figure 17). However, the CA2CA3 illustrated clear age-relationship differences between the techniques, potentially because of biases due to its small subfield size. For all the other subfields, our results were identical with HIPS or MAGeT techniques.

3.5. ASHS segmentation analysis

Hippocampal volumes extracted from ASHS and MAGeT segmentation techniques were investigated with age. Overall, ASHS volumes were significantly larger in the CA1, CA4DG and subiculum subfields and smaller in the CA2CA3 subfields. However, in all the subfields, no significant difference was found in the age-relationship of the volumes extracted from MAGeT and ASHS techniques (supplementary Figure 18).

4. Discussion

The aim of this paper was to examine the hippocampal subfields and WM subregions throughout healthy aging. Normalized for ICV, we

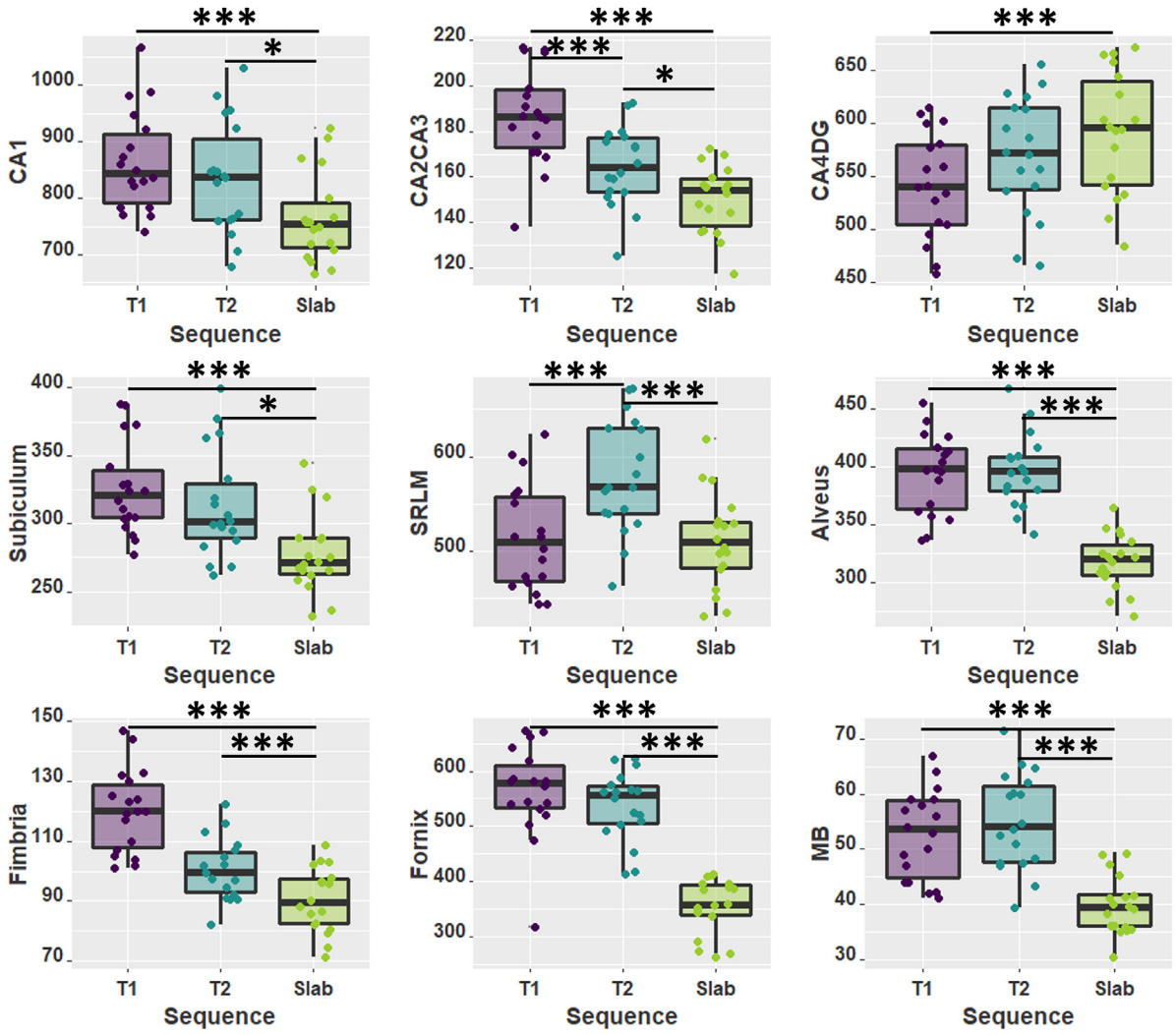

Fig. 6. Boxplots illustrating the right volume estimates from T1w, T2w and slab sequences from the same participants as well as the dependent 2-group Wilcoxon signed rank test results (Supplementary Table 1). Similar results were found in the left hemisphere (Supplementary figure 14). * $p < 0.05$; ** $p < 0.01$ and *** $p < 0.001$ after Bonferroni correction for 54 comparisons (18 subfields x 3 sequence types).

found that all subfields and WM subregions expressed a volumetric decrease with age, with the exception of CA1, subiculum, alveus and MB. These findings are in contradiction with previous investigations that demonstrated a significant impact of age on the CA1 (Daugherty et al., 2016; R. de Flores et al., 2015; Shing et al., 2011; Wisse et al., 2014; Wolf et al., 2015). Although, other studies found no volumetric change with age for the CA1 (Voineskos et al., 2015) or even an increase with age for the CA1 and alveus when accounting for ipsilateral hippocampal volume (Amaral et al., 2018). No clear age-related relationship between the CA1–3 volumes with age was demonstrated in another study, with a significant decrease in the body but not in the head and tail (Malykhin et al., 2017). Regarding WM subregions, the alveus, fimbria and fornix have been previously mostly used as anatomical landmarks for the GM subfields definition (La Joie et al., 2010; Malykhin et al., 2017; Mueller et al., 2007; Wisse et al., 2012). In the present study, WM subregions were considered as an integral part of the hippocampal circuitry. CA1 and alveus were found to be stable with age which reproduces, using a larger sample, previous findings from our group (Amaral et al., 2018).

Researchers have consistently shown the CA1 subfield to be the first hippocampal subfield to be impacted in AD (Adler et al., 2018; Amaral et al., 2018; de Flores, La Joie, and Chételat, 2015; Frisoni et al., 2008). Further, post-mortem studies of AD individuals have also found hippocampal amyloid and fibrillary tangles to firstly accumulate in CA1 (Braak and Braak, 1991; Braak et al., 2006). Therefore, we suggest that CA1 preservation in healthy aging could be of interest to identify early changes in hippocampus volume trajectories. SRLM has also been found to be especially impacted in patients with AD (Adler et al., 2018) as well as to a greater extent in *APOE4* carriers (Kerchner et al., 2014). Thus, it is interesting that our results also indicate a strong age-related atrophy in this subfield, although we did not find any *APOE4* effect. This could be in part because we did not have *APOE4* genotyping for a subset of our participants, preventing us from having enough data to fully investigate this genetic effect.

Complementary analyses were performed using ipsilateral hippocampal GM or WM volumes normalization. To our knowledge, this is the first study that analyzed the subfield volumetric relationship with respect to age while including global hippocampal atrophy. SRLM and fornix subregions expressed the highest relative volumetric impairment while the CA1 presented a relative volumetric preservation of its volumes with age. These findings replicate the results that we found using ICV normalization and provided a reasonable indication that the CA1 is a relatively preserved hippocampal subfield in healthy aging. Dissimilarities with other publications could be explained by the variability of the hippocampal subfield atlases definition used in the literature (Yushkevich et al., 2015). Indeed, we have previously found similar in-

	T1 vs T2		T1 vs SLAB		T2 vs SLAB	
STRUCTURES	LEFT	RIGHT	LEFT	RIGHT	LEFT	RIGHT
CA1	0.42	0.61	0.60	0.80	0.48	0.61
CA2CA3	0.66	0.66	0.70	0.79	0.87	0.75
CA4DG	0.65	0.80	0.87	0.88	0.71	0.75
SUB	0.76	0.70	0.70	0.72	0.64	0.72
SRLM	0.55	0.68	0.64	0.81	0.42	0.70
ALV	0.49	0.51	0.59	0.49	0.60	0.64
FIMB	0.35	0.72	0.64	0.77	0.52	0.80
FNX	0.36	0.47	0.53	0.60	0.28	0.52
MB	0.78	0.75	0.72	0.71	0.70	0.66

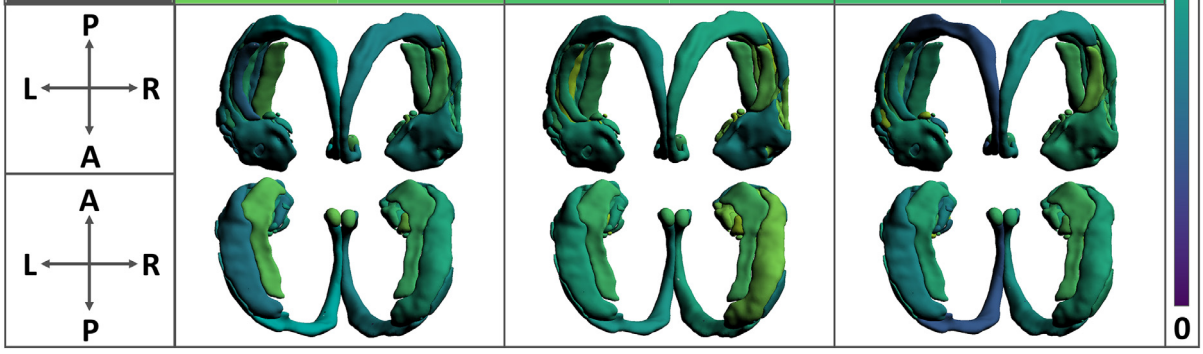

Fig. 7. Representation of ICC consistency (3,1) of hippocampal subfields volume estimates from T1w, T2w, and slab images. Colour scale serves as an indicator of the ICC values: yellow for an ICC of 1 and dark blue for an ICC of 0.

creases of the CA1 volumes (Amaral et al., 2018) after normalization using the same atlases and segmentation protocol in another cohort.

We hypothesize that the variability in findings within the literature may also be a function of the methodology used. For example, some groups approximated the hippocampal volumes using solely three contiguous slices (Daugherty et al., 2016) while others used nine slices (La Joie et al., 2010). Other studies have also expressed concern regarding the probable undersegmentation of CA1 and oversegmentation of the CA2–3 with FreeSurfer 5.3 (de Flores, La Joie, and Chételat, 2015). Also, only a few studies previously examined the WM subregions of the hippocampus (Amaral et al., 2018; Malykhin et al., 2017; Voineskos et al., 2015), potentially because slab sequences do not always provide proper field-of-view to study the entire hippocampal circuitry.

Also, in the present study, we included participants aged from 18 to 93, an age range which is larger than what is used in most papers studying the effect of age in the hippocampus subfields (Dounavi et al., 2020; Mueller et al., 2007; Wolf et al., 2015). Almost all the significant relationships with age demonstrated third order relationship, often with an inflection point close to 60 years old. Similarly, a critical age in the acceleration of hippocampal degeneration was previously identified at 63 years old (Yang et al., 2013).

Further, we investigated if our results were driven by a CNR relationship with age, since another study found a significant decrease in the hippocampal subfields with age (Knight et al., 2016). However, no age relationship of the CNR was found in our datasets which comforted us in our finding that the age effects that we identified were not related to a loss of contrast. However, while we defined our masks following a manual segmentation procedure previously used (Kulaga-Yoskovitz et al., 2015; Winterburn et al., 2013; Wood et al., 2015), others (Knight et al., 2016) have defined their WM masks in the SRLM. Thus, these protocol differences may explain the discrepancy regarding the CNR relationship with age in the hippocampus. Further, more investigation in regard to the age effect on the CNR would be of interest since we noticed a subtle downward trend in some hippocampal subfields using T2w images, particularly in light of the difference in ROIs used to examine the CNR.

A recent study compared different techniques to measure hippocampal subfield volumes and found that slab images were more sensitive to amyloid deposition and mild cognitive impairment status than whole brain T1w images (Mueller et al., 2018). Also, slab images have demonstrated highly consistent results with those from T1w images and slightly better detection of group effects in atrophy rates between patients with cognitive impairment and controls (Das et al., 2012). In our study, we found that volumes extracted from slab images demonstrated different age-related relationships than the volumes extracted from whole brain T1w and T2w images. This could be explained by the relatively small number of participants having a slab image in our study. Indeed, 64% of our original slab scans had to be excluded due to high motion artifact or incomplete coverage of hippocampal GM and WM subregions (Mueller et al., 2018) compared to 43% for T1w and 25% for T2w (Supplementary figure 2). Of note, it is unclear how consistent quality control has been applied in other slab volumetric studies.

To our knowledge, our paper used, for the first time, a high-resolution isotropic whole brain T2w sequence with isotropic voxel dimensions of 0.64 mm. The main advantage is that this acquisition provides a voxel volume of 0.26 mm^3 compared to 0.32 mm^3 for the commonly used slab sequence or 1 mm^3 for T1w sequence. Additionally, it allows for a clear delineation of the SRLM due to the hypo-intense contrast obtained in T2 contrast (Dounavi et al., 2020; Iglesias et al., 2015). This is of importance since it provides well-defined internal anatomical landmarks for segmentation protocols. T2w results primarily demonstrated similar results to T1w volume estimates with the supplementary benefit of demonstrating higher precision of segmentation according to visual inspection of the labels.

The results presented in this paper should be interpreted with respect to several considerations and limitations. First, an important consideration is that this study aims to draw a global conclusion in how the different hippocampal subfields evolve across the lifespan using cross-sectional data. Thus, even though longitudinal analysis would be essential and more appropriate to assess the “true” relationship of the hippocampal subfield volumes and age, these studies require tremendous resources, time and dedication from the participants. Therefore, in the context of our research, since no longitudinal study used various sequences on healthy participants to study the hippocampus subfields, we used cross-sectional studies to approximate and study this topic.

Secondly, to create this study, we used datasets from multiple sites. This led to different inclusion criteria with regards to how each site defined healthy participants (Supplementary methods). Also, participants were scanned in different scanners, especially in the ADNI dataset, which itself included multiple sites. To counterbalance this issue, we performed rigorous preprocessing, quality control (Bedford et al., 2020), and statistical analyses to standardize the quality and the intensity of the images included in this study to the best of our ability.

In addition, the differing results with the slab scans must be interpreted carefully since this dataset was limited to 116 participants. Unfortunately, we did not have access to slab data from healthy participants across the entire lifespan. That is why we decided to include the Test-retest dataset in order to increase the age-range of the slab sequence analyses. Thus, complementary analyses including more participants with a full coverage across the adult lifespan for the three modalities would be beneficial to validate these findings.

A significant challenge in the MRI subfield literature is the validation or derivation of the underlying segmentation protocols based on histological segmentation. To this end, few groups have successfully developed protocols that incorporate histology due to the scarcity of tissue and the significant challenge of histology-to-MRI nonlinear registration. The few groups that have derived their atlases from a histologically informed delineation, have to some extent used different strategies. For example, DeKraker et al., (2018) has relied heavily on the work from Ding and Van Hoesen (2015) while the work from Adler et al. (2014) relies more heavily on the Duvernoy atlas (2013). Thus, within the field there is still a lingering lack of clarity with respect to what "histologically-informed" methodologies to use. While there are still ongoing attempts to reconcile these disparities through the development of a harmonized subfield protocol that is informed, in some way, by the consensus segmentation of histological data (Olsen et al., 2019; Wisse et al., 2017), this remains a work in progress.

Nonetheless, it is important for us to acknowledge that, in the present paper, proper validation of the correspondence to ground-truth histology was unfortunately not possible. However, the atlases used were based on histological boundaries and rigorous quality control was completed to only include segmentations following these atlases histological rules (Amaral et al., 2018; Winterburn et al., 2013). Besides, in the current paper, we have examined the interplay between the acquisition technique and two additional segmentation algorithms that have been developed precisely for hippocampal subfield studies (Romero et al., 2017; Yushkevich et al., 2015). All of these techniques, including our own (Pipitone et al., 2014), which have been previously validated against manual segmentation.

Further, regarding the quality of segmentation, MAGeT has been developed to create a large template library given a much smaller sized input atlas library and then uses this template library in basic multi-atlas segmentation. The critical step here is that choosing 21 templates [optimal template number as demonstrated in (Pipitone et al., 2014)], led to having 105 atlas-to-template labels before doing a label fusion with the majority vote technique. Concretely, even if up to 52 atlas-to-template labels were giving erroneous information, these errors would not be propagated, because more than half (>53) would give correct information, which reduces tremendously the probability of propagating an error. Also, as reported in (Bhagwat et al., 2016), there is minimal impact in the quality of the segmentation according to the label fusion technique choice.

Another limitation relevant to the generalization of our findings is that our study is to a considerable extent dependent on our segmentation protocol (Chakravarty et al., 2013; Pipitone et al., 2014). Although MAGeT Brain has been validated by several studies (Makowski et al., 2018; Pipitone et al., 2014) and widely used to study the hippocampal subfields (Patel et al., 2020; Tardif et al., 2018; Voineskos et al., 2015), it remains that our subfield definitions are one among the multitude of atlases used in the literature. Supplementary results comparing MAGeT and HIPS (Manjon et al., 2020; Romero et al., 2017) segmentation techniques demonstrated almost similar relationships with age, highlighting that the relative preservation of the left CA1 with age for example was not due to a bias coming from MAGeT technique. Nonetheless, since HIPS uses the same atlases as MAGeT (Winterburn et al., 2013), here we solely investigated the impact of the segmentation technique choice but we did not investigate the impact of different subfield definitions. Thus, we also investigated ASHS segmentation technique which uses slab-specific anisotropic atlases as inputs. We did not find any difference in the volumetric age-relationships between MAGeT and ASHS outputs. However, as expected, overall volume differences were found between ASHS and MAGeT outputs, due to their different hippocampal subfield definitions. This highlights the importance of the work done by the Hippocampal Subfields Group (http://www.hippocampalsubfields.com/), which aims to standardize the definition of the medial temporal lobe segmentation, and to establish guidelines for subfield boundaries based on reference atlases and neuroanatomical landmarks visible postmortem and on MRI (Olsen et al., 2019; Yushkevich et al., 2015). A critical follow-up study would be to repeat this study to compare the effect of the MRI sequence using the complete and validated segmentation protocol described by the Hippocampal Subfield Group.

Outside of the sequence resolution and acquisition parameters, other parameters can impact segmentation quality. Of importance, the signal to noise ratio (SNR), contrast to noise ratio (CNR) and partial volume effect (PVE) are critical measures. Specifically, small regions (<200 mm^3) such as CA2CA3, fimbria or MB subregions or regions that are not very thick relative to the voxel size, such as the SRLM, may be prone to the effects of PVE. The extent of the impact of PVE on our final results remains unclear, and thus interpretations of age-related trends should be done with caution. Nonetheless, the work by Kerchner and colleagues (2010) has noted that the SRLM adjacent to the CA1 as being sensitive to the ageing process, which is inline with our findings. Additionally, other types of reliability measures, such as sharpness (Fonov and Collins, 2018; Shaw et al., 2019), would be of interest in future work.

A recently published commentary has highlighted concerns regarding the use of 1 mm^3 isotropic T1w scans for hippocampal subfield volumes studies (Wisse et al., 2020). While this work is important and demonstrates the need to consider carefully the advantages and drawbacks of using T1w 1 mm^3 images, a balanced discussion on other sequence benefits and flaws have not been reviewed. In light of our findings, we would like to advocate that a careful examination of the trade-offs between the different acquisition types should be done.

Another limitation of this paper is that we only used volumetric measurements to study the hippocampus. Nonetheless, other techniques could be used such as morphological methods which have previously provided spatially localized information about age-related hippocampal modifications (Voineskos et al., 2015; Yang et al., 2013). Another approach using hippocampal thickness, curvature, inner and outer surface textures and gyrification, has proven subfield-specific AP differences (DeKraker et al., 2020). The main advantage of morphological analyses is that it is independent from subfield definitions, which could be of interest while a consensus definition of the subfields is attained and published by the Hippocampal subfield group or to look at more local and subtle age-related changes.

For researchers interested in volumetric relationships with age, we advocate the use of large age-range datasets since we demonstrated that most of our age-related relationships were non-linear with an inflection point occurring at approximately age 60. Slab anisotropic scans seemed to find different age-effects compared to those found with T1w and T2w isotropic images. Also, because of the characteristics of slab scans, we would suggest to be especially careful in the volume approximation of small structures, such as CA2CA3, fornix or MB, since these structures can be locally thinner than 2 mm. Moreover, we advise researchers using slab datasets to keep in mind that it is likely that in general the subfield volumes are smaller compared to the estimation from standard T1w images (Das et al., 2012). Also, while slab images have the great advantage of having the highest in-plane resolution where some of the most

complicated subfield anatomical variation may occur, it clearly does so at the sacrifice of lower resolution in the anterior-to-posterior direction. Furthermore, the resulting voxel volume for the slab images used would be 0.32 mm^3 compared to 0.26 mm^3 for the T2w isotropic images we propose here. Finally, previous work from our own group demonstrated that the length of the long-axis is an important characteristic in the relationship of the hippocampus with age (Voineskos et al., 2015). Other groups have demonstrated a more highly curved hippocampus along the long axis in relation to age (Yang et al., 2013). In this context, we believe that the hippocampal shape modifications with age will cause issues at the level of resolving subfield borders that will increasingly "fall through the cracks" of the out of plane resolution limitations of the slab acquisition. To this end, it is critical to better examine the trade-offs between the different acquisition types carefully. Further, while investigating the accuracy of the segmentation protocols that we used (MAGeT, HIPS and ASHS) in the context of a "curving" hippocampus with age would be of interest, assessing the hippocampal shape would require a completely separate set of methods and analyses (such as voxel-based morphometry, deformation-based morphometry, or MAGeT surface-based morphometry measures for example), which unfortunately would be out of the scope of this study.

Finally, for new scanning protocols, we encourage researchers to consider using high-resolution isotropic T2w images, which have shown promising sensitivity and accuracy. T2w images also demonstrated smaller volumes in most of the subfields compared to T1w images, but we predict that they may better estimate the "true" volumes since they have higher isotropic resolution and contrast in key areas. T2w images rendered the same age-related relationships as T1w in most structures, although stronger age-related relationships were identified for the CA1 when compared to T1w images. Furthermore, high-resolution T2w images have demonstrated high-quality segmentation from visual inspection compared to slab and T1w scans (Fig. 1), and provide whole brain images that permit hippocampal WM subregion investigation.

To conclude, we compared a wide range of datasets attempting to elucidate the relationship between hippocampal subfields with age. Certain subfields appeared to show reliable and reproducible relationships with age across various datasets. Nonetheless, the sensitivity of slab images to age-related changes in the hippocampus subfields deserves further exploratory analysis.

Acknowledgments

A. Bussy receives support from the Alzheimer Society of Canada. M. Chakravarty is funded by the Weston Brain Institute, the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada and Fondation de Recherches Santé Québec.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12–2–0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.;Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.;Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's 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.

Data collection and sharing for this project was provided by the Cambridge Centre for Ageing and Neuroscience (CamCAN). CamCAN funding was provided by the UK Biotechnology and Biological Sciences Research Council (grant number BB/H008217/1), together with support from the UK Medical Research Council and University of Cambridge, UK.

We would like to thank José V. Manjón (IBIME, UPV, Spain) and Pierrick Coupé (LaBRI UMR 5800, Université de Bordeaux, CNRS, France) for the access to the VolBrain pipeline and computing allocations for the HIPS segmentations presented in this manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.117931.

References

Adler, D.H., Pluta, J., Kadivar, S., Craige, C., Gee, J.C., Avants, B.B., Yushkevich, P.A., 2014. Histology-derived volumetric annotation of the human hippocampal subfields in postmortem MRI. Neuroimage 84, 505–523.

Adler, D.H., Wisse, L.E.M., Ittyerah, R., Pluta, J.B., Ding, S.-L., Xie, L., Wang, J., Kadivar, S., Robinson, J.L., Schuck, T., Trojanowski, J.Q., Grossman, M., Detre, J.A., Elliott, M.A., Toledo, J.B., Liu, W., Pickup, S., Miller, M.I., Das, S.R., Yushkevich, P.A., 2018. Characterizing the human hippocampus in aging and Alzheimer's disease using a computational atlas derived from ex vivo MRI and histology. Proc. Natl. Acad. Sci. U.S.A. 115 (16), 4252–4257.

Akaike, H., 1974. A new look at the statistical model identification. In: Springer Series in Statistics, pp. 215–222. doi:10.1007/978-1-4612-1694-0_16.

Amaral, R.S.C., Park, M.T.M., Devenyi, G.A., Lynn, V., Pipitone, J., Winterburn, J., Chavez, S., Schira, M., Lobaugh, N.J., Voineskos, A.N., Pruessner, J.C., Chakravarty, M.M.Alzheimer's Disease Neuroimaging Initiative, 2018. Manual segmentation of the fornix, fimbria, and alveus on high-resolution 3T MRI: application via fully-automated mapping of the human memory circuit white and grey matter in healthy and pathological aging. Neuroimage 170, 132–150.

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 12 (1), 26–41.

Bedford, S.A., Park, M.T.M., Devenyi, G.A., Tullo, S., Germann, J., Patel, R., Anagnostou, E., Baron-Cohen, S., Bullmore, E.T., Chura, L.R., Craig, M.C., Ecker, C., Floris, D.L., Holt, R.J., Lenroot, R., Lerch, J.P., Lombardo, M.V., Murphy, D.G.M., Raznahan, A., Chakravarty, M.M., 2020. Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. Mol. Psychiatry 25 (3), 614–628.

Bellon, E.M., Haacke, E.M., Coleman, P.E., Sacco, D.C., Steiger, D.A., Gangarosa, R.E., 1986. MR artifacts: a review. AJR Am. J. Roentgenol. 147 (6), 1271–1281.

Bender, A.R., Keresztes, A., Bodammer, N.C., Shing, Y.L., Werkle-Bergner, M., Daugherty, A.M., Yu, Q., Kühn, S., Lindenberger, U., Raz, N., 2018. Optimization and validation of automated hippocampal subfield segmentation across the lifespan. Hum. Brain Mapp. 39 (2), 916–931.

Bhagwat, N., Pipitone, J., Winterburn, J.L., Guo, T., Duerden, E.G., Voineskos, A.N., Lepage, M., Miller, S.P., Pruessner, J.C., Chakravarty, M.M., 2016. Manual-protocol inspired technique for improving automated MR Image segmentation during label fusion. Front. Neurosci. 10, 325.

Bobinski, M., de Leon, M.J., Wegiel, J., Desanti, S., Convit, A., Saint Louis, L.A., Rusinek, H., Wisniewski, H.M., 2000. The histological validation of post mortem magnetic resonance imaging-determined hippocampal volume in Alzheimer's disease. Neuroscience 95 (3), 721–725.

Braak, H., Alafuzoff, I., Arzberger, T., Kretzschmar, H., Del Tredici, K., 2006. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. Acta Neuropathol. 112 (4), 389–404.

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. In Acta Neuropathol. (Vol. 82, Issue 4, pp. 239–259). 10.1007/bf00308809

Bussy, A., Snider, B.J., Coble, D., Xiong, C., Fagan, A.M., Cruchaga, C., Benzinger, T.L.S., Gordon, B.A., Hassenstab, J., Bateman, R.J., Morris, J.C., 2019. Effect of apolipoprotein E4 on clinical, neuroimaging, and biomarker measures in noncarrier participants in the Dominantly Inherited Alzheimer Network. Neurobiol. Aging 75, 42–50.

Campbell, S., MacQueen, G., 2004. The role of the hippocampus in the pathophysiology of major depression. J. Psychiatry Neurosci.: JPN. https://psycnet.apa.org/record/2005-02137-002.

Chakravarty, M.M., Sadikot, A.F., Germann, J., Bertrand, G., Collins, D.L., 2008. Towards a validation of atlas warping techniques. Med. Image Anal. 12 (6), 713–726.

Chakravarty, M.M., Sadikot, A.F., Germann, J., Hellier, P., Bertrand, G., Collins, D.L., 2009. Comparison of piece-wise linear, linear, and nonlinear atlas-to-patient warping

techniques: analysis of the labeling of subcortical nuclei for functional neurosurgical applications. Hum. Brain Mapp. 30 (11), 3574–3595.
Chakravarty, M.M., Steadman, P., van Eede, M.C., Calcott, R.D., Gu, V., Shaw, P., Raznahan, A., Collins, D.L., Lerch, J.P., 2013. Performing label-fusion-based segmentation using multiple automatically generated templates. Hum. Brain Mapp. 34 (10), 2635–2654.
Cherbuin, N., Leach, L.S., Christensen, H., Anstey, K.J., 2007. Neuroimaging and APOE genotype: a systematic qualitative review. Dement. Geriatr. Cogn. Disord. 24 (5), 348–362.
Chételat, G., Fouquet, M., Kalpouzos, G., Denghien, I., De la Sayette, V., Viader, F., Mézenge, F., Landeau, B., Baron, J.C., Eustache, 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 (6), 1721–1731.
Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J. Comput. Assist. Tomogr. 18 (2), 192–205.
Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L., Pericak-Vance, M.A., 1993. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261 (5123), 921–923.
Coupé, P., Catheline, G., Lanuza, E., Manjón, J.V., 2017. Towards a unified analysis of brain maturation and aging across the entire lifespan: a MRI analysis. Hum. Brain Mapp. 38 (11), 5501–5518.
Dadar, M., Fonov, V.S., Collins, D.L.Alzheimer's Disease Neuroimaging Initiative, 2018. A comparison of publicly available linear MRI stereotaxic registration techniques. Neuroimage 174, 191–200.
Das, S.R., Avants, B.B., Pluta, J., Wang, H., Suh, J.W., Weiner, M.W., Mueller, S.G., Yushkevich, P.A., 2012. Measuring longitudinal change in the hippocampal formation from in vivo high-resolution T2-weighted MRI. Neuroimage 60 (2), 1266–1279.
Daugherty, A.M., Bender, A.R., Raz, N., Ofen, N., 2016. Age differences in hippocampal subfield volumes from childhood to late adulthood. Hippocampus 26 (2), 220–228.
de Flores, R., La Joie, R., Chételat, G., 2015a. Structural imaging of hippocampal subfields in healthy aging and Alzheimer's disease. Neuroscience 309, 29–50.
de Flores, R., La Joie, R., Landeau, B., Perrotin, A., Mézenge, F., de La Sayette, V., Eustache, F., Desgranges, B., Chételat, G., 2015b. Effects of age and Alzheimer's disease on hippocampal subfields: comparison between manual and FreeSurfer volumetry. Hum. Brain Mapp. 36 (2), 463–474.
DeKraker, J., Ferko, K.M., Lau, J.C., Köhler, S., Khan, A.R., 2018. Unfolding the hippocampus: an intrinsic coordinate system for subfield segmentations and quantitative mapping. Neuroimage 167, 408–418.
DeKraker, J., Lau, J.C., Ferko, K.M., Khan, A.R., Köhler, S., 2020. Hippocampal subfields revealed through unfolding and unsupervised clustering of laminar and morphological features in 3D BigBrain. Neuroimage 206, 116328.
Derix, J., Yang, S., Lüsebrink, F., Fiederer, L.D.J., Schulze-Bonhage, A., Aertsen, A., Speck, O., & Ball, T. (2014). Visualization of the amygdalo-hippocampal border and its structural variability by 7T and 3T magnetic resonance imaging. In Hum. Brain Mapp. (Vol. 35, Issue 9, pp. 4316–4329). 10.1002/hbm.22477
Ding, S.-.L., Van Hoesen, G.W., 2015. Organization and detailed parcellation of human hippocampal head and body regions based on a combined analysis of cyto-and chemoarchitecture. J. Comp. Neurol. 523 (15), 2233–2253.
Dounavi, M.-.E., Mak, E., Wells, K., Ritchie, K., Ritchie, C.W., Su, L., O' Brien, J.T., 2020. Volumetric alterations in the hippocampal subfields of subjects at increased risk of dementia. Neurobiol. Aging doi:10.1016/j.neurobiolaging.2020.03.006.
Duvernoy, H.M., 2005. The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections With MRI. Springer Science & Business Media.
Duvernoy, H.M., 2013. The Human Hippocampus: An Atlas of Applied Anatomy. J.F. Bergmann-Verlag.
Dwork, A.J. (1997). Postmortem studies of the hippocampal formation in schizophrenia. In Schizophr Bull (Vol. 23, Issue 3, pp. 385–402). 10.1093/schbul/23.3.385
Eriksson, S.H., Thom, M., Bartlett, P.A., Symms, M.R., McEvoy, A.W., Sisodiya, S.M., Duncan, J.S., 2008. PROPELLER MRI visualizes detailed pathology of hippocampal sclerosis. Epilepsia 49 (1), 33–39.
Eskildsen, S.F., Coupé, P., Fonov, V., Manjón, J.V., Leung, K.K., Guizard, N., Wassef, S.N., Østergaard, L.R., Collins, D.L.Alzheimer's Disease Neuroimaging Initiative, 2012. BEaST: brain extraction based on nonlocal segmentation technique. Neuroimage 59 (3), 2362–2373.
Fonov, V.S., & Louis Collins, D. (2018). Comparison of different methods for average anatomical templates creation: do we really gain anything from a diffeomorphic framework? (p. 277087). 10.1101/277087
Frisoni, G.B., Ganzola, R., Canu, E., Rüb, U., Pizzini, F.B., Alessandrini, F., Zoccatelli, G., Beltramello, A., Caltagirone, C., Thompson, P.M., 2008. Mapping local hippocampal changes in Alzheimer's disease and normal ageing with MRI at 3 Tesla. Brain: J. Neurol. 131 (Pt 12), 3266–3276.
Gonneaud, J., Arenaza-Urquijo, E.M., Fouquet, M., Perrotin, A., Fradin, S., de La Sayette, V., Eustache, F., Chételat, G., 2016. Relative effect of APOE $\varepsilon 4$ on neuroimaging biomarker changes across the lifespan. Neurology 87 (16), 1696–1703.
Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N., Friston, K.J., Frackowiak, R.S., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14 (1 Pt 1), 21–36.
Goubran, M., Crukley, C., de Ribaupierre, S., Peters, T.M., Khan, A.R., 2013. Image registration of ex-vivo MRI to sparsely sectioned histology of hippocampal and neocortical temporal lobe specimens. Neuroimage 83, 770–781.
Goubran, M., Rudko, D.A., Santyr, B., Gati, J., Szekeres, T., Peters, T.M., Khan, A.R., 2014. In vivo normative atlas of the hippocampal subfields using multi-echo susceptibility imaging at 7 Tesla. Hum. Brain Mapp. 35 (8), 3588–3601.
Habes, M., Toledo, J.B., Resnick, S.M., Doshi, J., Van der Auwera, S., Erus, G., Janowitz, D., Hegenscheid, K., Homuth, G., Völzke, H., Hoffmann, W., Grabe, H.J., Davatzikos, C., 2016. Relationship between APOE genotype and structural MRI measures throughout adulthood in the study of health in Pomerania population-based cohort. AJNR Am. J. Neuroradiol. 37 (9), 1636–1642.
Heckers, S., 2001. Neuroimaging studies of the hippocampus in schizophrenia. Hippocampus 11 (5), 520–528.
Iglesias, J.E., Augustinack, J.C., Nguyen, K., Player, C.M., Player, A., Wright, M., Roy, N., Frosch, M.P., McKee, A.C., Wald, L.L., Fischl, B., Van Leemput, K.Alzheimer's Disease Neuroimaging Initiative, 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.
Iosifescu, D.V., Shenton, M.E., Warfield, S.K., Kikinis, R., Dengler, J., Jolesz, F.A., McCarley, R.W., 1997. An automated registration algorithm for measuring MRI subcortical brain structures. Neuroimage 6 (1), 13–25.
Jack Jr, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L. Whitwell, J., Ward, C., 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging: Offic. J. Int. Soc. Magn. Reson.Med. 27 (4), 685–691.
Kerchner, G.A., Berdnik, D., Shen, J.C., Bernstein, J.D., Fenesy, M.C., Deutsch, G.K., Wyss–Coray, T., Rutt, B.K., 2014. APOE $\varepsilon 4$ worsens hippocampal CA1 apical neuropil atrophy and episodic memory. Neurology 82 (8), 691–697.
Kerchner, G.A., Hess, C.P., Hammond-Rosenbluth, K.E., Xu, D., Rabinovici, G.D., Kelley, D.A.C., Vigneron, D.B., Nelson, S.J., Miller, B.L, 2010. Hippocampal CA1 apical neuropil atrophy in mild Alzheimer disease visualized with 7-T MRI. Neurology 75 (15), 1381–1387.
Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.-.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 46 (3), 786–802.
Knight, M.J., McCann, B., Tsivos, D., Couthard, E., Kauppinen, R.A., 2016. Quantitative T1 and T2 MRI signal characteristics in the human brain: different patterns of MR contrasts in normal ageing. MAGMA 29 (6), 833–842.
Kulaga-Yoskovitz, J., Bernhardt, B.C., Hong, S.-.J., Mansi, T., Liang, K.E., van der Kouwe, A.J.W., Smallwood, J., Bernasconi, A., Bernasconi, N., 2015. Multi-contrast submillimetric 3 Tesla hippocampal subfield segmentation protocol and dataset. Sci. Data 2, 150059.
Kurth, F., Cherbuin, N., Luders, E., 2017. The impact of aging on subregions of the hippocampal complex in healthy adults. Neuroimage 163, 296–300.
Laakso, M.P., Frisoni, G.B., Könönen, M., Mikkonen, M., Beltramello, A., Geroldi, C., Bianchetti, A., Trabucchi, M., Soininen, H., Aronen, H.J., 2000. Hippocampus and entorhinal cortex in frontotemporal dementia and Alzheimer's disease: a morphometric MRI study. Biol. Psychiatry 47 (12), 1056–1063.
La Joie, R., Fouquet, M., Mézenge, F., Landeau, B., Villain, N., Mevel, K., Pélerin, A., Eustache, F., Desgranges, B., Chételat, G., 2010. Differential effect of age on hippocampal subfields assessed using a new high-resolution 3T MR sequence. Neuroimage 53 (2), 506–514.
Makowski, C., Béland, S., Kostopoulos, P., Bhagwat, N., Devenyi, G.A., Malla, A.K., Joober, R., Lepage, M., Chakravarty, M.M., 2018. Evaluating accuracy of striatal, pallidal, and thalamic segmentation methods: comparing automated approaches to manual delineation. Neuroimage 170, 182–198.
Malykhin, N.V., Carter, R., Seres, P., Coupland, N.J., 2010. Structural changes in the hippocampus in major depressive disorder: contributions of disease and treatment. J. Psychiatry Neurosci.: JPN 35 (5), 337–343.
Malykhin, N.V., Huang, Y., Hrybouski, S., Olsen, F., 2017. Differential vulnerability of hippocampal subfields and anteroposterior hippocampal subregions in healthy cognitive aging. Neurobiol. Aging 59, 121–134.
Manjon, J.V., Romero, J.E., & Coupe, P. (2020). DeepHIPS: a novel Deep Learning based Hippocampus Subfield Segmentation method. In arXiv [q-bio.QM]. arXiv. http://arxiv.org/abs/2001.11789
Marques, J.P., Norris, D.G., 2018. How to choose the right MR sequence for your research question at 7 T and above? Neuroimage 168, 119–140.
Mazerolle, M., 2006. Improving data analysis in herpetology: using Akaike's Information Criterion (AIC) to assess the strength of biological hypotheses. Amphibia-Reptilia: Publication of the Societas Europaea Herpetologica 27 (2), 169–180.
Mueller, S.G., Stables, L., Du, A.T., Schuff, N., Truran, D., Cashdollar, N., Weiner, M.W., 2007. Measurement of hippocampal subfields and age-related changes with high resolution MRI at 4T. Neurobiol. Aging 28 (5), 719–726.
Mueller, S.G., Weiner, M.W., 2009. Selective effect of age, Apo e4, and Alzheimer's disease on hippocampal subfields. Hippocampus 19 (6), 558–564.
Mueller, S.G., Yushkevich, P.A., Das, S., Wang, L., Van Leemput, K., Iglesias, J.E., Alpert, K., Mezher, A., Ng, P., Paz, K., & Weiner, M.W. (2018). Systematic comparison of different techniques to measure hippocampal subfield volumes in ADNI2. In NeuroImage: Clinical (Vol. 17, pp. 1006–1018). 10.1016/j.nicl.2017.12.036
Mugler III, J.P., Brookeman, J.R, 1990. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). Magn. Reson. Med.: Offic.J. Soc. Magn. Reson. Medicine/Soc. Magn. Reson. Med. 15 (1), 152–157.
Muñoz-Ruiz, M.Á., Hartikainen, P., Koikkalainen, J., Wolz, R., Julkunen, V., Niskanen, E., Herukka, S.-.K., Kivipelto, M., Vanninen, R., Rueckert, D., Liu, Y., Lötjönen, J., Soininen, H., 2012. Structural MRI in frontotemporal dementia: comparisons between hippocampal volumetry, tensor-based morphometry and voxel-based morphometry. PloS One 7 (12), e52531.
Nelson, M.D., Saykin, A.J., Flashman, L.A., Riordan, H.J., 1997. Hippocampal volume reduction in schizophrenia as assessed by magnetic resonance imaging: a meta-analytic study. Schizophr. Res. 24 (1–2), 153. doi:10.1016/s0920-9964(97)82438-3, Issues.

Olsen, R.K., Carr, V.A., Daugherty, A.M., La Joie, R., Amaral, R.S.C., Amunts, K., Augustinack, J.C., Bakker, A., Bender, A.R., Berron, D., Boccardi, M., Bocchetta, M., Burggren, A.C., Chakravarty, M.M., Chételat, G., de Flores, R., DeKraker, J., Ding, S.-L., Geerlings, M.I.Hippocampal Subfields Group, 2019. Progress update from the hippocampal subfields group. Alzheimer's Dement.: J. Alzheimer's Assoc. 11, 439–449.

Palombo, D.J., Amaral, R.S.C., Olsen, R.K., Müller, D.J., Todd, R.M., Anderson, A.K., Levine, B, 2013. KIBRA polymorphism is associated with individual differences in hippocampal subregions: evidence from anatomical segmentation using high-resolution MRI. J. Neurosci. 33 (32), 13088–13093.

Parekh, M.B., Rutt, B.K., Purcell, R., Chen, Y., Zeineh, M.M., 2015. Ultra-high resolution in-vivo 7.0T structural imaging of the human hippocampus reveals the endfolial pathway. Neuroimage 112, 1–6.

Patel, R., Steele, C.J., Chen, A.G.X., Patel, S., Devenyi, G.A., Germann, J., Tardif, C.L., Chakravarty, M.M, 2020. Investigating microstructural variation in the human hippocampus using non-negative matrix factorization. Neuroimage 207, 116348.

Pereira, J.B., Valls-Pedret, C., Ros, E., Palacios, E., Falcón, C., Bargalló, N., Bartrés-Faz, D., Wahlund, L.-.O., Westman, E., Junque, C., 2014. Regional vulnerability of hippocampal subfields to aging measured by structural and diffusion MRI. Hippocampus 24 (4), 403–414.

Pipitone, J., Park, M.T.M., Winterburn, J., Lett, T.A., Lerch, J.P., Pruessner, J.C., Lepage, M., Voineskos, A.N., Chakravarty, M.M.Alzheimer's Disease Neuroimaging Initiative, 2014. Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. Neuroimage 101, 494–512.

Pluta, J., Yushkevich, P., Das, S., Wolk, D., 2012. In vivo analysis of hippocampal subfield atrophy in mild cognitive impairment via semi-automatic segmentation of T2-weighted MRI. J. Alzheimer's Dis.: JAD 31 (1), 85–99.

Pol, L.A., van de, van de Pol, L.A., Hensel, A., Barkhof, F., Gertz, H.J., Scheltens, P., van der Flier, W.M., 2006. Hippocampal atrophy in Alzheimer disease: age matters. Neurology 66 (2), 236–238. doi:10.1212/01.wnl.0000194240.47892.4d.

Rawle, M.J., Davis, D., Bendayan, R., Wong, A., Kuh, D., Richards, M., 2018. Apolipoprotein-E (Apoe) ϵ4 and cognitive decline over the adult life course. Transl. Psychiatry, 8 doi:10.1038/s41398-017-0064-8.

Raz, N., Rodrigue, K.M., Head, D., Kennedy, K.M., Acker, J.D., 2004. Differential aging of the medial temporal lobe: a study of a five-year change. Neurology 62 (3), 433–438.

Reuter, M., Tisdall, M.D., Qureshi, A., Buckner, R.L., van der Kouwe, A.J.W., Fischl, B., 2015. Head motion during MRI acquisition reduces grey matter volume and thickness estimates. Neuroimage 107, 107–115.

Romero, J.E., Coupé, P., Manjón, J.V., 2017. HIPS: a new hippocampus subfield segmentation method. Neuroimage 163, 286–295.

Shafto, M.A., Tyler, L.K., Dixon, M., Taylor, J.R., Rowe, J.B., Cusack, R., Calder, A.J., Marslen-Wilson, W.D., Duncan, J., Dalgleish, T., Henson, R.N., Brayne, C., Matthews, F.E.Cam-CAN, 2014. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. BMC Neurol. 14, 204.

Shaw, T.B., Bollmann, S., Atcheson, N.T., Guo, C., Fripp, J., Salvado, O., Barth, M., 2019. Non-linear realignment improves hippocampus subfield segmentation reliability. Neuroimage doi:10.1101/597856.

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

Smith, T.B., Nayak, K.S., 2010. MRI artifacts and correction strategies. Imaging Med. https://www.openaccessjournals.com/articles/mri-artifacts-and-correction-strategies-11010.html.

Stockmeier, C.A., Mahajan, G.J., Konick, L.C., Overholser, J.C., Jurjus, G.J., Meltzer, H.Y., Uylings, H.B.M., Friedman, L., Rajkowska, G, 2004. Cellular changes in the postmortem hippocampus in major depression. Biol. Psychiatry 56 (9), 640–650.

Sullivan, E.V., Marsh, L., Mathalon, D.H., Lim, K.O., Pfefferbaum, A., 1995. Age-related decline in MRI volumes of temporal lobe grey matter but not hippocampus. Neurobiol. Aging 16 (4), 591–606.

Sullivan, E.V., Marsh, L., Pfefferbaum, A., 2005. Preservation of hippocampal volume throughout adulthood in healthy men and women. Neurobiol. Aging 26 (7), 1093–1098.

Svarer, C., Madsen, K., Hasselbalch, S.G., Pinborg, L.H., Haugbøl, S., Frøkjaer, V.G., Holm, S., Paulson, O.B., Knudsen, G.M., 2005. MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. Neuroimage 24 (4), 969–979.

Tardif, C.L., Devenyi, G.A., Amaral, R.S.C., Pelleieux, S., Poirier, J., Rosa-Neto, P., Breitner, J., Chakravarty, M.M.PREVENT-AD Research Group, 2018. Regionally specific changes in the hippocampal circuitry accompany progression of cerebrospinal fluid biomarkers in preclinical Alzheimer's disease. Hum. Brain Mapp. 39 (2), 971–984.

Taylor, J.R., Williams, N., Cusack, R., Auer, T., Shafto, M.A., Dixon, M., Tyler, L.K., Cam-CAN, Henson, R.N, 2017. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. Neuroimage 144, 262–269. doi:10.1016/j.neuroimage.2015.09.018.

Thomas, D.L., De Vita, E., Roberts, S., Turner, R., Yousry, T.A., Ordidge, R.J., 2004. High-resolution fast spin echo imaging of the human brain at 4.7 T: implementation and sequence characteristics. Magn. Reson. Med.: Offic. J. Soc. Magn. Reson. Med./Soc. Magn. Reson. Med. 51 (6), 1254–1264.

Tullo, S., Patel, R., Devenyi, G.A., Salaciak, A., Bedford, S.A., Farzin, S., Wlodarski, N., Tardif, C.L., Group, P.-A.R., Breitner, J.C.S., et al., 2019. MR-based age-related effects on the striatum, globus pallidus, and thalamus in healthy individuals across the adult lifespan. Hum. Brain Mapp. 40 (18), 5269–5288.

Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: improved N3 bias correction. IEEE Trans. Med. Imaging 29 (6), 1310–1320.

Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L.L., Augustinack, J., Dickerson, B.C., Golland, P., Fischl, B., 2009. Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI. Hippocampus 19 (6), 549–557. doi:10.1002/hipo.20615.

Voineskos, A.N., Winterburn, J.L., Felsky, D., Pipitone, J., Rajji, T.K., Mulsant, B.H., Chakravarty, M.M., 2015. Hippocampal (subfield) volume and shape in relation to cognitive performance across the adult lifespan. Hum. Brain Mapp. 36 (8), 3020–3037.

Winterburn, J.L., Pruessner, J.C., Chavez, S., Schira, M.M., Lobaugh, N.J., Voineskos, A.N., Chakravarty, M.M., 2013. A novel in vivo atlas of human hippocampal subfields using high-resolution 3 T magnetic resonance imaging. Neuroimage 74, 254–265.

Winterburn, J., Pruessner, J.C., Sofia, C., Schira, M.M., Lobaugh, N.J., Voineskos, A.N., Chakravarty, M.M., 2015. High-resolution in vivo manual segmentation protocol for human hippocampal subfields using 3T magnetic resonance imaging. J. Vis. Exp. 105, e51861.

Wisse, L.E.M., Biessels, G.J., Heringa, S.M., Kuijf, H.J., Koek, D.H.L., Luijten, P.R., Geerlings, M.I.Utrecht Vascular Cognitive Impairment (VCI) Study Group, 2014. Hippocampal subfield volumes at 7T in early Alzheimer's disease and normal aging. Neurobiol. Aging 35 (9), 2039–2045.

Wisse, L.E.M., Chételat, G., Daugherty, A.M., Flores, R., Joie, R., Mueller, S.G., Stark, C.E.L., Wang, L., Yushkevich, P.A., Berron, D., Raz, N., Bakker, A., Olsen, R.K., Carr, V.A, 2020. Hippocampal subfield volumetry from structural isotropic 1 mm 3 MRI scans: a note of caution. Hum. Brain Mapp. doi:10.1002/hbm.25234.

Wisse, L.E.M., Daugherty, A.M., Olsen, R.K., Berron, D., Carr, V.A., Stark, C.E.L., Amaral, R.S.C., Amunts, K., Augustinack, J.C., Bender, A.R., Bernstein, J.D., Boccardi, M., Bocchetta, M., Burggren, A., Chakravarty, M.M., Chupin, M., Ekstrom, A., de Flores, R., Insausti, R.Hippocampal Subfields Group, 2017. A harmonized segmentation protocol for hippocampal and parahippocampal subregions: why do we need one and what are the key goals? Hippocampus 27 (1), 3–11.

Wisse, L.E.M., Gerritsen, L., Zwanenburg, J.J.M., Kuijf, H.J., Luijten, P.R., Biessels, G.J., Geerlings, M.I, 2012. Subfields of the hippocampal formation at 7 T MRI: in vivo volumetric assessment. Neuroimage 61 (4), 1043–1049.

Wolf, D., Fischer, F.U., de Flores, R., Chételat, G., Fellgiebel, A., 2015. Differential associations of age with volume and microstructure of hippocampal subfields in healthy older adults. Hum. Brain Mapp. 36 (10), 3819–3831.

Wood, B., Knight, M.J., Tsivos, D., Oliver, R., Coulthard, E., Kauppinen, R.A., 2015. Magnetic resonance scanning and image segmentation procedure at 3 T for volumetry of human hippocampal subfields. Biomed. Spectrosc. Imaging 4 (2), 197–208.

Yang, X., Goh, A., Chen, S.-H.A., Qiu, A., 2013. Evolution of hippocampal shapes across the human lifespan. Hum. Brain Mapp. 34 (11), 3075–3085.

Yushkevich, P.A., Amaral, R.S.C., Augustinack, J.C., Bender, A.R., Bernstein, J.D., Boccardi, M., Bocchetta, M., Burggren, A.C., Carr, V.A., Chakravarty, M.M., Chételat, G., Daugherty, A.M., Davachi, L., Ding, S.-.L., Ekstrom, A., Geerlings, M.I., Hassan, A., Huang, Y., Iglesias, J.E.Hippocampal Subfields Group (HSG), 2015a. 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.

Yushkevich, P.A., Wang, H., Pluta, J., Das, S.R., Craige, C., Avants, B.B., Weiner, M.W., Mueller, S., 2010. Nearly automatic segmentation of hippocampal subfields in in vivo focal T2-weighted MRI. Neuroimage 53 (4), 1208–1224.

Zhao, W., Wang, X., Yin, C., He, M., Li, S., Han, Y., 2019. Trajectories of the Hippocampal Subfields Atrophy in the Alzheimer's Disease: a Structural Imaging Study. Front. Neuroinform. 13, 13.

Zheng, F., Cui, D., Zhang, L., Zhang, S., Zhao, Y., Liu, X., Liu, C., Li, Z., Zhang, D., Shi, L., Liu, Z., Hou, K., Lu, W., Yin, T., Qiu, J., 2018. The volume of hippocampal subfields in relation to decline of memory recall across the adult lifespan. Front. Aging Neurosci. 10, 320.