

# Hippocampal volume across age: Nomograms derived from over 19,700 people in UK Biobank

Lisa Nobis<sup>a,\*</sup>, Sanjay G. Manohar<sup>b,c</sup>, Stephen M. Smith<sup>d,e</sup>, Fidel Alfaro-Almagro<sup>d,e</sup>, Mark Jenkinson<sup>d,e</sup>, Clare E. Mackay<sup>a,e</sup>, Masud Husain<sup>b,c,d</sup>

<sup>a</sup> Oxford Centre for Human Brain Activity, Wellcome Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, Oxford, UK

<sup>b</sup> Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

<sup>c</sup> Department of Experimental Psychology, University of Oxford, Oxford, UK

<sup>d</sup> Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

<sup>e</sup> Oxford Health NHS Foundation Trust, Oxford, UK

## ARTICLE INFO

### Keywords:

Atrophy  
Brain ageing  
MRI  
Normative values  
Sex differences

## ABSTRACT

Measurement of hippocampal volume has proven useful to diagnose and track progression in several brain disorders, most notably in Alzheimer's disease (AD). For example, an objective evaluation of a patient's hippocampal volume status may provide important information that can assist diagnosis or risk stratification of AD. However, clinicians and researchers require access to age-related normative percentiles to reliably categorise a patient's hippocampal volume as being pathologically small. Here we analysed effects of age, sex, and hemisphere on the hippocampus and neighbouring temporal lobe volumes, in 19,793 generally healthy participants in the UK Biobank. A key finding of the current study is a significant acceleration in the rate of hippocampal volume loss in middle age, more pronounced in females than in males. In this report, we provide normative values for hippocampal and total grey matter volume as a function of age for reference in clinical and research settings. These normative values may be used in combination with our online, automated percentile estimation tool to provide a rapid, objective evaluation of an individual's hippocampal volume status. The data provide a large-scale normative database to facilitate easy age-adjusted determination of where an individual hippocampal and temporal lobe volume lies within the normal distribution.

## 1. Introduction

One of the most common brain imaging markers used in clinical research to study severity and progression of Alzheimer's Disease (AD) is hippocampal volume on a structural MRI scan (Frisoni et al., 2010; Ahmed et al., 2014). Both longitudinal and cross-sectional studies have reported reduced volume of the hippocampus in patients with AD or mild cognitive impairment (MCI) compared to healthy controls (Henneman et al., 2009; Shi et al., 2009; Frankó et al., 2013). In addition, a meta-analysis of nine studies found a 3.33% difference in atrophy rate between AD and controls (Barnes et al., 2009). Increased atrophy of the hippocampus has also been associated with neurofibrillary tangle and amyloid plaque deposition, which are considered to be the hallmark features of Alzheimer's disease (Kril et al., 2002; Schuff et al., 2009). Similarly, using the rate of hippocampal atrophy, researchers have been able to distinguish between those with MCI who progressed to AD, and those who did not (Frankó et al., 2013).

Importantly, the regions of the hippocampus with the highest atrophy rate also presented with the most severe amyloid deposition (Frankó et al., 2013). Hence, hippocampal volume is considered to be a useful – and widely available – proxy to measure disease burden and progression in AD, including in clinical trials (Mielke et al., 2012; Frankó et al., 2013; Kishi et al., 2015; Choe et al., 2016).

Precise estimations of hippocampal volume may also facilitate earlier diagnosis, with some researchers reporting that the rate of hippocampal atrophy deviated from a healthy trajectory about 5.5 years before the stage of clinical AD was reached (Chételat et al., 2005). Indeed, it appears that baseline hippocampal volume and atrophy rate may be more suitable than whole brain volume to distinguish between MCI and controls (Henneman et al., 2009).

It is now appreciated that estimation of hippocampal volume may also be useful for measuring disease burden or progression in other disorders. For example, hippocampal atrophy has been related to impaired episodic memory in patients with multiple sclerosis (Koenig

\* Corresponding author at: Oxford Centre for Human Brain Activity, Department of Psychiatry, Warneford Ln, Oxford OX3 7JX, UK.

E-mail address: [Lisa.nobis@psych.ox.ac.uk](mailto:Lisa.nobis@psych.ox.ac.uk) (L. Nobis).

<https://doi.org/10.1016/j.nicl.2019.101904>

Received 12 March 2019; Received in revised form 13 June 2019

Available online 19 June 2019

2213-1582/ © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

et al., 2014) and temporal lobe epilepsy (Reyes et al., 2018), and can predict cognitive impairment in patients with Parkinson's disease (Kandiah et al., 2014). Similarly, a cross-sectional study in patients with autoimmune encephalitis (Chakos et al., 2005) reported that the severity of the disease course is associated with lower hippocampal volume, which also predicted worse cognitive outcome (Finke et al., 2017).

In psychiatric disorders too, hippocampal volume has been implicated as an important imaging correlate. For example, reduced hippocampal volume has been repeatedly linked to depression (Campbell et al., 2004; Videbech and Ravnkilde, 2004), and a meta-analysis has reported that smaller hippocampi in depressed patients predict lower response rates to antidepressant drugs (Colle et al., 2016). Similarly, in schizophrenia, hippocampal volume is reduced in chronic cases compared to healthy controls (Adriano et al., 2012). In addition, in a recent study on the effects of cognitive remediation therapy for cognitive impairment in schizophrenia, improvement was correlated with increased hippocampal volume post-treatment (Morimoto et al., 2018). These findings demonstrate that hippocampal volume estimations have the potential to provide important information to assist diagnosis, risk stratification and possibly even monitor the effects of intervention. However, several key challenges in the processing, analysis, and interpretation of structural MRI scans outside of a research setting have so far prevented this method from being used in standard clinical practice.

First, in order to calculate hippocampal volume from a structural MRI scan, the image has to be processed with software requiring expert knowledge. Automated brain segmentation tools, e.g. the FIRST tool in FMRIB Software Library (FSL) (Jenkinson and Smith, 2001; Jenkinson et al., 2002), or FreeSurfer (<http://surfer.nmr.mgh.harvard.edu/>), and openly available standardised processing pipelines (Alfaro-Almagro et al., 2018), may therefore improve the accessibility of this method in clinical practice.

Second, hippocampal atrophy has not only been described in pathological, but also in healthy ageing. A meta-analysis estimated the average yearly rate of hippocampal atrophy at 1.4% in healthy ageing, and 4.7% in AD (Barnes et al., 2009). Thus, in order to reliably categorise a patient's hippocampal volume, clinicians and researchers require access to age-related normative percentiles to distinguish between healthy and pathological ageing. Few studies have attempted to provide such normative values – or nomograms – not least because of methodological obstacles. The acquisition of large MRI datasets is costly, time consuming, and involves highly specialised hardware and staff. Previously published studies on hippocampal volume in the general population are therefore often limited by small sample sizes, low statistical power, and cohort effects (Ioannidis, 2011; Button et al., 2013; Fraser et al., 2015; Nord et al., 2017). In addition, other potentially confounding factors, such as method used, sex, head size, years of education, smoking status, and effects of normal ageing have to be considered when applying such normative percentiles.

Here we report a first attempt to provide detailed normative information on the hippocampus and neighbouring temporal lobe volumes in relation to age, sex, total grey matter volume, symmetry, and lifestyle factors in 19,793 generally healthy older adults in the UK Biobank resource (<http://www.ukbiobank.ac.uk>). We chose a model-free sliding window approach to study the relationships of temporal lobe volumes and total grey matter with age, as this method makes few assumptions and does not impose a linear relationship. While efforts to harmonize analysis across MRI datasets have not yet been successful, the large dataset analysed here allows an attempt to provide robust normative data when these are used with estimates obtained using the same, or similar, acquisition protocols, hardware and pre-processing pipelines.

We also supply a webtool, where a patient's estimated brain volume may be entered to calculate their hippocampal volume status as compared to age-matched norm values ([www.win.ox.ac.uk/open-neuroimaging/analyse-your-data/hippocampal-volume-tool](http://www.win.ox.ac.uk/open-neuroimaging/analyse-your-data/hippocampal-volume-tool)). These

norms will become especially useful with the acquisition of longitudinal health outcomes of participants in UK Biobank. The UK Biobank cohort includes over 500,000 participants aged 40–69 years, of which a subset of 100,000 participants will undergo brain imaging. Within this subset, 6000 participants are likely to have developed AD by 2027 (Miller et al., 2016). UK Biobank has been granted access to the UK National Health Service records and can thus follow up on future health outcomes of participants, making it a powerful resource to study disease progression. In addition, both the acquisition and analysis of MRI are standardised across the scanning sites according to publicly available protocols designed by the UK Biobank Imaging Working Group ([www.ukbiobank.ac.uk/expert-working-groups](http://www.ukbiobank.ac.uk/expert-working-groups)). This increases generalisability of results and resolves some of the issues of previous studies using different scanners and protocols.

## 2. Methods

### 2.1. Participants

The most recent release of brain imaging data of the UK Biobank includes scans of 20,542 participants aged 45–80 years at the time of scanning. Participants with self-reported neurological and psychiatric disorders, substance abuse disorders or a history of head trauma were excluded from our analyses, leaving a sample of 19,793 generally healthy participants scanned in 2014–2018.

### 2.2. MRI acquisition and analysis

Image acquisition and pre-processing of MRI scans have been performed by UK Biobank. Brain images were acquired on a Siemens Skyra 3.0T scanner (Siemens Medical Solutions, Germany) with a 32-channel head coil. T1-weighted images with 1 mm<sup>3</sup> isotropic resolution were previously analysed with FMRIB Software Library (FSL) (<http://fsl.fmrib.ox.ac.uk/fsl>), with image-derived phenotypes (IDPs – imaging summary statistics such as brain volume and hippocampal volume) made available for general access. More detailed information on MRI acquisition and analysis have been reported elsewhere (Miller et al., 2016; Alfaro-Almagro et al., 2018). UK Biobank also published a standardised MRI analysis pipeline (FMRIB's Biobank Pipeline version 1.0) that is freely available to the public, including the source code ([https://git.fmrib.ox.ac.uk/falmagro/UK\\_biobank\\_pipeline\\_v1](https://git.fmrib.ox.ac.uk/falmagro/UK_biobank_pipeline_v1)) (Alfaro-Almagro et al., 2018).

### 2.3. Statistics

All calculations were performed in MATLAB 2017b with the exception of Joinpoint regression, which was computed with the Joinpoint Regression Program (version 4.6.0.0; Surveillance Research Program, National Cancer Institute) (Kim et al., 2000).

#### 2.3.1. Data preparation

Outliers in brain volume estimations were identified and excluded based on a Median Absolute Deviation (MAD) denominated distance from the median of > 5, as this method is more robust to non-Gaussian distributions. The cut-off of 5 was chosen by visual inspection of data histograms. Brain volumes were always corrected for scanning date, and when indicated, for head size and age. This was done with a General Linear Model, 'regressing out' variance accounted for by the confounds. Brain volumes were corrected for scanning date to adjust for gradual changes in the scanner hardware, an effect called scanner drift. These changes lead to variation of the signal distribution in scans over time, resulting in over- or underestimations of volumes (Takao et al., 2011). Head size was corrected for using a head scaling variable obtained from the MRI scan, which is the volumetric scaling for the transformation of the native head image to standard space, with final scaling being driven by outer-skull surface (Smith et al., 2002, 2004). A number of studies have shown that global and regional brain volumes

correlate positively with head size (Barnes et al., 2010). Thus, when investigating sex differences in brain volume, head size should be corrected for. This was done by regression rather than using a ratio measure (e.g. brain volume/head size), as not all brain volumes scale proportionately with head size (Barnes et al., 2010). However, measures of head size related variables vary widely in the literature (e.g. total intracranial volume, total brain volume, MRI head scaling, or body height), leading to heterogeneous results. We therefore report both head size corrected and un-corrected results.

### 2.3.2. Analysis

For comparison we also present analyses for volumes of some structures near the hippocampus within the temporal lobe. Available cortical grey matter volume IDPs (in mm<sup>3</sup>) within the UK Biobank dataset that are reported here include superior temporal gyrus, middle temporal gyrus, inferior temporal gyrus, fusiform gyrus, parahippocampal gyrus and temporal pole. These were previously calculated using FSL-FAST in conjunction with cortical atlas regions-of-interest (Miller et al., 2016; Alfaro-Almagro et al., 2018).

t-tests (with Bonferroni-correction for resulting p-values) were calculated for sex differences in age and all included volumes of interest (unpaired t-test), as well as for differences in hemisphere (paired t-test), hypertension status (unpaired t-test), smoking status (unpaired t-test, for hippocampus and total grey matter only) and hippocampal volume estimations by method (FSL-FIRST, vs. FSL-FAST within Atlas-based regions of interest; paired t-test). While FSL-FIRST is a segmentation/registration tool that estimates subcortical volumes based on shape and appearance models (Patenaude et al., 2011), FSL-FAST is a brain tissue segmentation tool (Zhang et al., 2001), with which subcortical volumes were estimated from grey matter within Atlas-based regions of interest.

Pearson correlations were calculated for associations of the brain volumes of interest with BMI. ANOVAs were run for the effects of level of education and smoking status on the brain volumes of interest. Education level was binarized into higher education degree (college/university degree, NVQ), and lower education degree (A levels, O levels, CSE). In addition to scanning date, brain volumes of interest were corrected for age and head size before analysing effects of sex, hypertension status, smoking status, education, and estimation method. No additional corrections were applied before testing for volume asymmetries.

For visualisation purposes, volume differences based on sex, hemisphere, hypertension status, and education are plotted as proportional differences, calculated with  $2 * \frac{(Mean\ volume(a) - Mean\ volume(b))}{(Mean\ volume(a) + Mean\ volume(b))}$ . This was done to allow for comparison across different brain areas with different sizes, as the proportional differences can be plotted on one scale. Effect sizes were calculated with the 'Measures of Effect Size (MES) Toolbox' in MATLAB (Hentschke and Stuetgen, 2011), using the Hedges' g statistic.

### 2.3.3. Normative percentiles

The 2.5th, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 97.5th percentiles for hippocampal volume (left and right) and total grey matter were calculated separately for males and females. This was done using residuals from a corresponding sliding-window analysis (see below), that were added back onto the mean of the brain volume estimations. These values were then used as the age-adjusted input to calculate the percentiles. Percentiles are provided for volumes corrected for head size and scanning date. Percentiles corrected for scanning date, but uncorrected for head size can be found in the Supplementary material.

### 2.3.4. Sliding-window curves for relationship of volume and age

To analyse trajectories and quantiles of measured volumes as a function of age, we applied a sliding-window, model-free analysis, in which an age-window of observations of fixed age-quantile width was moved along the age distribution (based on function 'conditionalPlot' in

<https://osf.io/vmabg/>; (Manohar, 2019)). The windows overlapped, so that each window contained 10% of the participants. The analysis of interest was performed for each window, providing in this case the mean volume of brain structure, such as hippocampal volume. The volumes were then smoothed using the moving average method with a gaussian kernel of 20. Results from using varying smoothing kernels and quantile widths can be found in the Supplementary Material. The mean brain volume was plotted against the mean bin centre of age, along with standard errors of the mean. This provided an estimate of mean brain volume that varied smoothly as a function of age. This method enables representative contours to be calculated for each quantile as a function of age, in a non-parametric, data-driven manner. Here, we present sliding-window curves across age for all available temporal lobe areas, and total grey matter volume. In addition, curves for the ratio of the individual temporal lobe volumes to the rest of grey matter volume are shown, as well as the slope of the hippocampal volume versus age (in Supplementary Material), using bootstrapped confidence intervals. Volumes were corrected for head size and scanning date. The trajectory for hippocampal volume across age, uncorrected for head size, is included in the Supplementary Material.

We tested for differences in the peak of the ratio curves between males and females. This was performed using a permuted t-statistic, correcting p-values for false discovery rate. Specifically, we permuted matched age bins of the male and female groups, i.e. shuffling males and females within age-windows. We then calculated the p-value as the proportion of permuted datasets ( $n = 5000$ ), which produced a mean difference between the peaks of the curves at least as extreme as the one observed from the actual data.

This data-driven analysis is particularly useful for large samples. As opposed to traditional curve fitting methods, it makes few assumptions, has only one free parameter (number of samples per bin), and improves visualisation and interpretation of results and corresponding errors. Note that as there are fewer participants at the minimum and maximum age limits compared to around the mean age, this method leads to wider age ranges for the lowest and highest percentile bins. As shown in Supplementary Fig. S14, similar effects to our main analysis were observed when using windows of fixed age bins (5 years width), rather than bins of fixed numbers of observations. Again, the mean brain volume was computed within each age bin, and plotted against the mean bin centre of that bin, along with standard errors.

As a parametric alternative to using sliding-window analysis for the estimation of normative percentiles, we present results from a generalised additive model for location, scale and shape (GAMLSS) (Stasinopoulos and Rigby, 2007) in Supplementary Material. This method has previously been used for growth chart percentiles published by the World Health Organization (World Health Organization Multicentre Growth Reference Study Group, 2006), and has some advantages over a non-parametric method as it produces smooth percentile curves, and allows extrapolation to the youngest and oldest age groups. However, the sliding-window analysis was chosen here as it generates a true representation of the raw data, makes no assumptions about the shape of the distribution, and has minimal free parameters.

### 2.3.5. Joinpoint regression

To determine if there were significant points at which the rate of hippocampal atrophy changed across age, we used joinpoint regression. This is a regression analysis recommended by the National Cancer Institute (NCI, <https://surveillance.cancer.gov/joinpoint/>, (Kim et al., 2000)) to estimate time points of change in trend data, often used for changes in cancer incidence rates. The analysis fits separate line segments to trend data, which are connected at positions called joinpoints. No joinpoints would suggest no changes in slope, and thus no changes in, for example, cancer incidence. Here we applied joinpoint regression to test for the presence of one or more joinpoints (i.e. inflection points) along the brain volume trajectories across age. Thus, instead of cancer incidence rate, here the dependent variable was mean volume per 1-

year, non-overlapping age window. Bins of 1-year width were chosen to retain adequate accuracy with a feasible number of datapoints. Similarly, instead of diagnosis year, the independent variable used here was age group.

The Joinpoint Regression Program fits between zero and a chosen maximum number of joinpoints to the data. Starting from zero joinpoints, we tested whether more joinpoints are statistically significant ( $p < .05$ ) using a Monte Carlo Permutation method, assuming uncorrelated errors. The maximum was chosen at 4 joinpoints, as recommended for our number of data points using grid search (Kim et al., 2000). This recommendation is based on allowing at least seven data points to find one joinpoint, and at least two data points between consecutive joinpoints. A joinpoint can also not occur within two data points from the beginning or end of a series.

This analysis was done with hippocampal volume, rest of total grey matter volume (total grey matter minus hippocampal volume), superior temporal gyrus, middle temporal gyrus, inferior temporal gyrus, fusiform gyrus, parahippocampal gyrus, and temporal pole volume, for males and females separately. All volumes were corrected for scanning date and head size. More detailed information on methods used within the Joinpoint Regression Program can be found at <https://surveillance.cancer.gov/help/joinpoint/>.

### 3. Results

#### 3.1. Demographics

Demographics as well as means and standard deviations for uncorrected hippocampal volume with respect to head size are summarised in Table 1. Number of observations exclude outliers (e.g. 137 outliers removed for left hippocampal volume, 134 outliers removed for right hippocampal volume).

#### 3.2. Method of segmentation affects hippocampal volume estimates

Mean hippocampal volumes differed between methods. On average estimates calculated with FSL-FAST with atlas were larger ( $mean = 4285.08 \text{ mm}^3$ ,  $std = 309.32 \text{ mm}^3$ ) than estimates calculated with FSL-FIRST ( $mean = 3863.40 \text{ mm}^3$ ,  $std = 353.61 \text{ mm}^3$ ,  $p < .001$ , Hedges'  $g = 1.27$ ). This overestimation of hippocampal volume with FSL-FAST compared to FSL-FIRST is further demonstrated in Fig. 1, which shows a Bland-Altman plot of average hippocampal volume. Estimates were corrected for head size, scanning date, and age. However, there were no differences in the overall results from the joinpoint and sliding-window analyses between the two hippocampal volume estimates (see sections below). As FSL-FIRST is the recommended and most widely used method for subcortical segmentation, we present analyses throughout using the IDPs that were generated using FSL-FIRST.

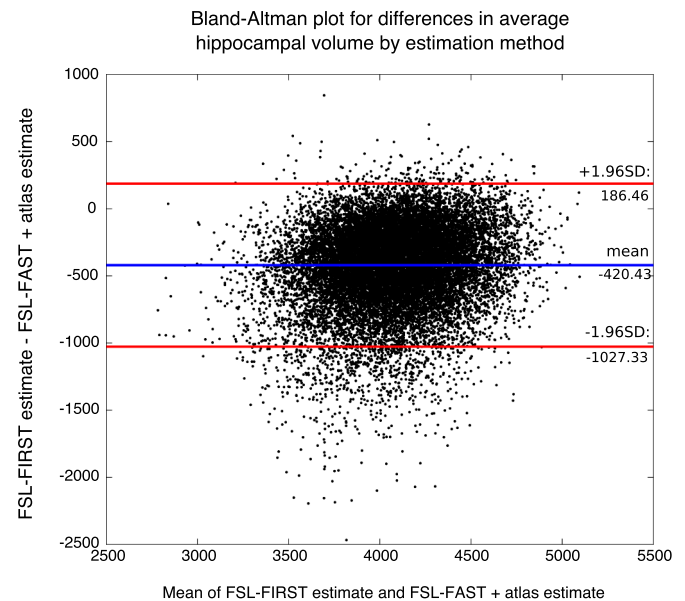
#### 3.3. Normative percentiles

Example nomograms for the 2.5th, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 97.5th percentiles for left and right hippocampal volume in

**Table 1**  
Demographics of sample.

	Age (Mean $\pm$ std)	Total grey matter in $\text{mm}^3$ (Mean $\pm$ std)	Average hippocampus in $\text{mm}^3$ (Mean $\pm$ std)	N
Female	62.31 $\pm$ 7.34 years	594,250 $\pm$ 48,028 $\text{mm}^3$	3,765.18 $\pm$ 366.75 $\text{mm}^3$	10,463
Male	63.67 $\pm$ 7.58 years	641,307 $\pm$ 51,379 $\text{mm}^3$	3,972.97 $\pm$ 431.03 $\text{mm}^3$	9330
Total	62.95 $\pm$ 7.48 years	616,484 $\pm$ 54,917 $\text{mm}^3$	3,863.94 $\pm$ 411.56 $\text{mm}^3$	19,793

Mean age was 62.95 years ( $std = 7.48$  years), with males slightly older than females ( $p < .001$ , Hedges'  $g = 0.18$ ). Mean total grey matter volume (uncorrected for head size) was significantly larger in males than females ( $p < .001$ , Hedges'  $g = 0.95$ ), as was the case for uncorrected average hippocampal volume ( $p < .001$ , Hedges'  $g = 0.52$ ).



**Fig. 1.** Bland-Altman plot for differences in average hippocampal volume by estimation method.

Mean average hippocampal volume of FSL-FIRST and FSL-FAST + atlas estimates on the x-axis is plotted against the difference between the two methods on the y-axis. The blue line represents the mean of the difference, red lines represent 1.96SD above and below the mean.

**Table 2**  
Head size correction parameters.

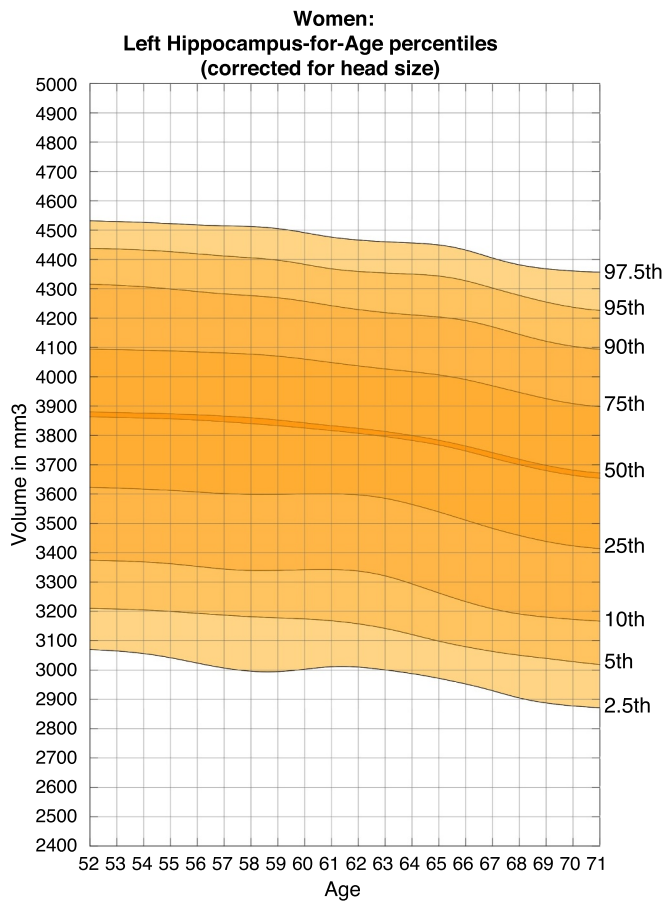
	Left hippocampus	Right hippocampus	Total grey matter
b	-1,340.89	-1,532.24	-354,545.84

$$\text{Volume}_{\text{corrected}} = \text{Volume}_{\text{uncorrected}} - [(\text{Head Scaling} - 1.29874) * b]$$

women and men, corrected for head size, are shown in Table 2, Figs. 2–5. Overlapping, age-adjusted windows of 10% of observations are plotted, with age on the x-axis corresponding to the median age of participants within the overlapping quantile windows. Because there are fewer observations within particularly young (starting at 45 years) and particularly old (up to 80 years) participants, quantile bins for the youngest and oldest participants are wider, and thus have a median age around 50 and 75 years, respectively. Additional moving average smoothing with a gaussian kernel of 20 resulted in percentiles plotted across ages 51 years to 72 years (52 years to 73 years for males).

The nomograms are provided separately for females and males as the previous analyses have shown differential sex by age trajectories. Additional nomograms split by hemisphere and sex for total grey matter, as well as plots with volumes corrected for head scaling, and plots for hippocampus-to-total grey matter ratio can be found in the Supplementary Material (Suppl. Figs. S1–12). A comparison of the sliding-window method to the GAMLSS method is shown in Supplementary Fig. S13.





**Fig. 2.** Nomogram of left hippocampal volume for females, corrected for head size.

Figures show the quantiles of hippocampal volume for the group of individuals in each age window. The x-axis indicates the median age of the window. The mid line indicates the median. For someone with a given age and volume, a percentile can be read off the chart, indicating the proportion of the Biobank cohort who have a hippocampal volume below that of the person.

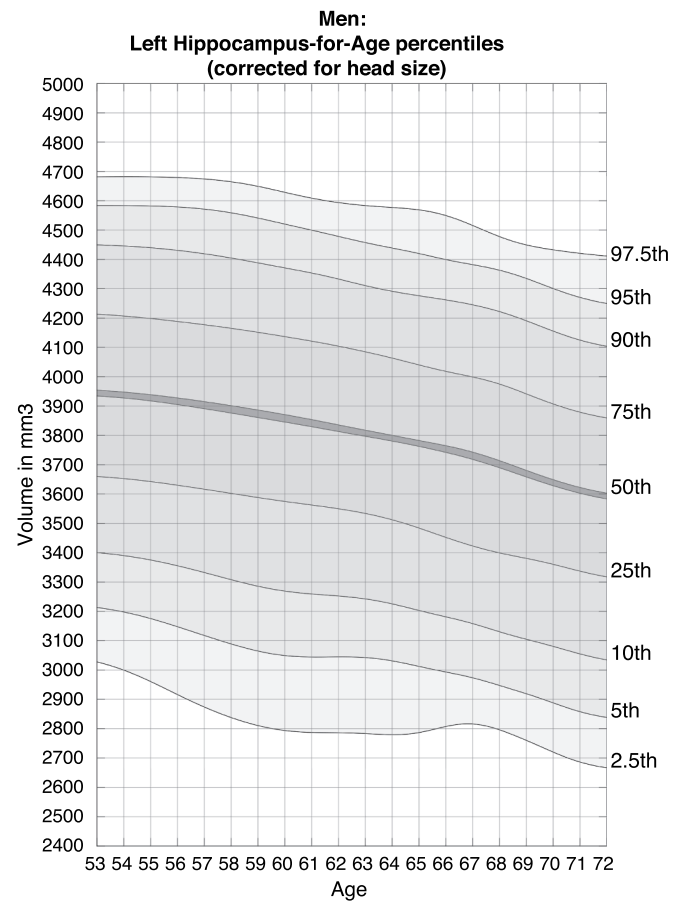
Head-size corrected percentiles can be used by applying the following formula to the corresponding uncorrected brain volume, using the ‘Head Scaling’ variable calculated from the automated pipeline, and given slopes for either left hippocampus, right hippocampus, or total grey matter. The formula is based on a general linear model and calculates head-size corrected volumes of left, right, or grey matter volume by subtracting the product of demeaned head scaling and the corresponding slope from the uncorrected volume.

The provided nomogram webtool ([www.win.ox.ac.uk/open-neuroimaging/analyse-your-data/hippocampal-volume-tool](http://www.win.ox.ac.uk/open-neuroimaging/analyse-your-data/hippocampal-volume-tool)) enables clinicians to quickly calculate a patient's percentile by entering the volume estimation from the automated analysis pipeline. The tool then automatically applies head size correction, if the option is selected. For example, a female aged 64 years with a left hippocampal volume of 3.15 cm<sup>3</sup> (corrected for head size) would score within the 5th percentile. This means that 95% of the women closest to her age in the current sample had a larger volume, highlighting potentially pathological atrophy and providing additional information for the diagnosing clinician.

### 3.4. Brain volume characteristics

#### 3.4.1. Temporal lobe volumes differed between hemispheres

A paired *t*-test resulted in a significant difference between right and left hippocampal volume (uncorrected for head size) with *right larger*



**Fig. 3.** Nomogram of left hippocampal volume for males, corrected for head size.

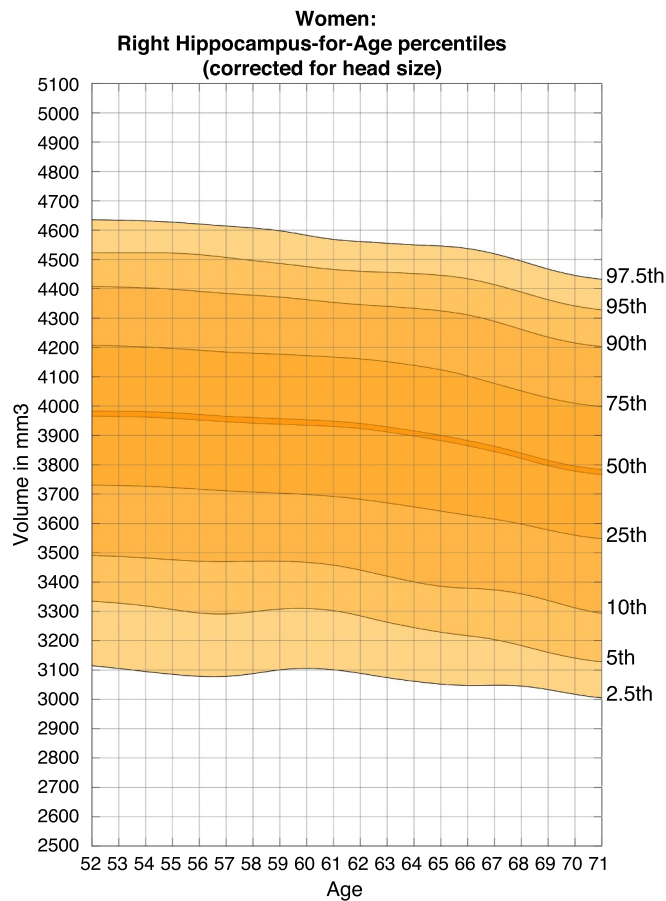
Figures show the quantiles of hippocampal volume for the group of individuals in each age window. The x-axis indicates the median age of the window. The mid line indicates the median. For someone with a given age and volume, a percentile can be read off the chart, indicating the proportion of the Biobank cohort who have a hippocampal volume below that of the person.

than left volume asymmetry (see Fig. 6 for percent difference,  $p < .001$ , *Hedges' G* = 0.24). The same effect was found in the superior temporal gyrus ( $p < .001$ , *Hedges' G* = 0.80) and inferior temporal gyrus ( $p < .001$ , *Hedges' G* = 0.12). On the other hand, there was a significant *left larger than right* volume asymmetry for the fusiform gyrus ( $p < .001$ , *Hedges' G* = 1.30) and parahippocampal gyrus ( $p < .001$ , *Hedges' G* = 0.48). Asymmetry effects of the same direction in the temporal pole were also significant, however with a very small effect size (*Hedges' G* = 0.04). There was no significant difference between left and right hemisphere for middle temporal gyrus ( $p = .61$ ). See Supplementary Table S4 for means, standard deviations, and significance tests for all temporal lobe volumes.

Head size was not corrected for as left/right differences are within-subject, and will be largely unaffected by head size. There was no correlation between handedness and asymmetry for any of the tested volumes (calculated by asymmetry = left volume - right volume; all  $p > \alpha/7 = 0.007$ ).

#### 3.4.2. Temporal lobe volumes by sex

Means, standard deviations, and significance tests for all brain volumes by sex, corrected for head size and age, can be found in Supplementary Table S5. There was no significant difference of mean (average) hippocampal volume between males and females ( $p > .05$ ). Sex differences for the remaining temporal lobe volumes, and for total grey matter volume were significant in different directions, though



**Fig. 4.** Nomogram of right hippocampal volume for females, corrected for head size.

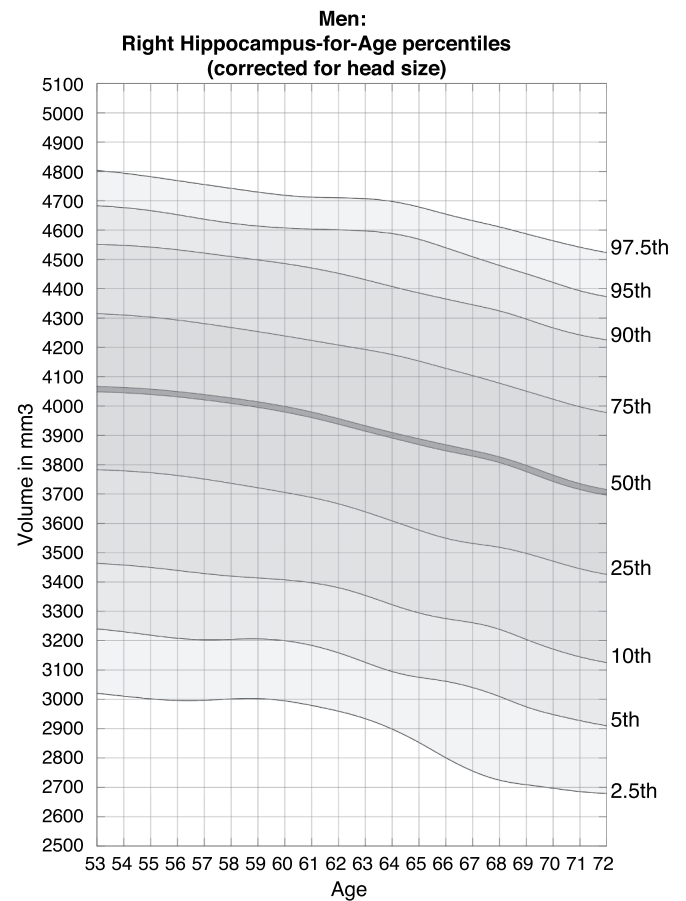
Figures show the quantiles of hippocampal volume for the group of individuals in each age window. The x-axis indicates the median age of the window. The mid line indicates the median. For someone with a given age and volume, a percentile can be read off the chart, indicating the proportion of the Biobank cohort who have a hippocampal volume below that of the person.

effect sizes were mostly very small (See Fig. 7 for percent differences). Small-to-medium effect sizes were seen in total grey matter volume (females larger than males), parahippocampal gyrus (males larger than females), and temporal pole (males larger than females).

#### 3.4.3. Trajectory of mean hippocampal volume with age demonstrate an inflection

We plot mean hippocampal volume, total grey matter volume (both corrected for head size), and hippocampal volume as a percentage of grey matter volume as a function of age. To plot this, we used a sliding-window method, using windows of 10% of the population (Fig. 8). The x-axis displays the median age of participants within the overlapping quantile windows. Because there are fewer observations within particularly young and particularly old participants, the median ages of the quantile bins for the youngest and oldest participants are around 50 and 75 years, respectively. Plots for left and right hippocampal volumes, bilateral hippocampal volume uncorrected for head size, as well as plots for fixed 5-year age-bins rather than fixed quantiles for bilateral hippocampal volume, are provided in the Supplementary Material (Suppl. Figs. S14–S15, S21).

Whereas whole brain volume appears to decline linearly as a function of age for both men and women (Fig. 8B), in women the rate of hippocampal volume loss increases at ~60 years (Fig. 8A). This inflection point (i.e. maximum slope), can be calculated by numerical differentiation, as shown in Supplementary Fig. S16 (females: bin(max



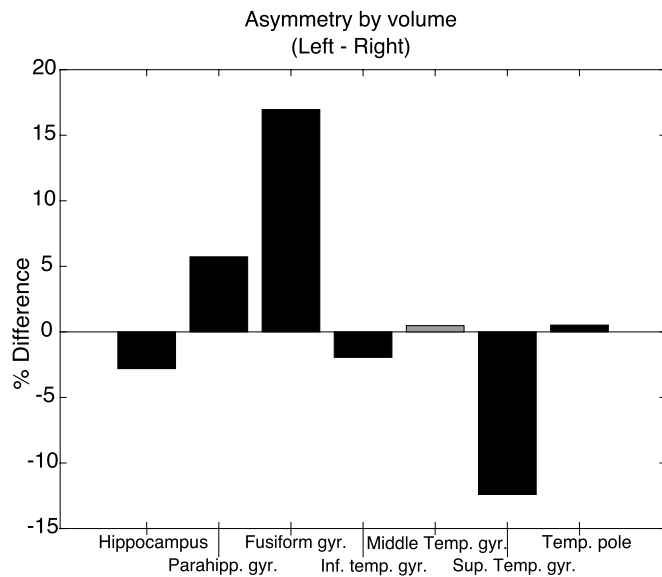
**Fig. 5.** Nomogram of right hippocampal volume for males, corrected for head size.

Figures show the quantiles of hippocampal volume for the group of individuals in each age window. The x-axis indicates the median age of the window. The mid line indicates the median. For someone with a given age and volume, a percentile can be read off the chart, indicating the proportion of the Biobank cohort who have a hippocampal volume below that of the person.

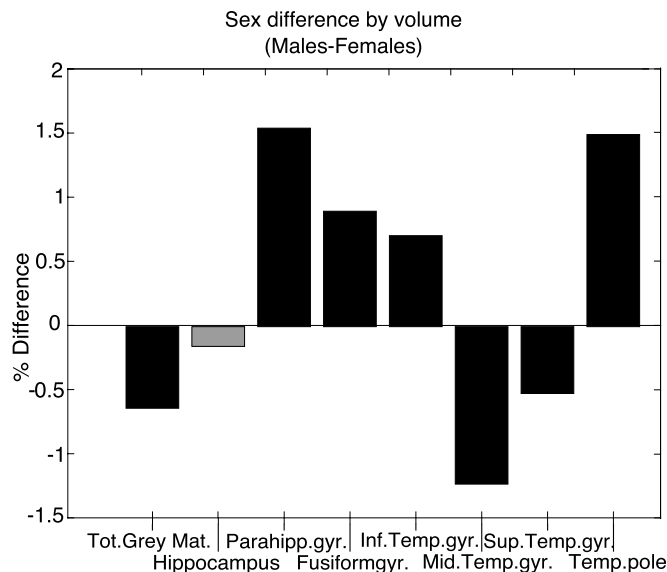
slope) = 58.67–61.28 years, max slope =  $-0.27$ ; males: bin(max slope) = 44.56–52.73 years, max slope =  $-2.52$ ).

To compare hippocampal volume more directly to the rest of the grey matter, the ratio between these two volumes was calculated. In females, the ratio of hippocampal volume to the rest of grey matter (Fig. 8C) peaks around age 67 years (bin(max ratio) = 66.33 years – 68.44 years), suggesting that the rate of volume loss in the hippocampus is smaller relative to the rest of grey matter. As age increases further, this pattern is then reversed. In males the ratio of hippocampal volume relative to the rest of grey matter peaks around age 63 years (bin(max ratio) = 61.69 years – 64.32 years), before declining thereafter. These sex differences in the age of transition (maximum ratio) were significant. Males reached peak hippocampal ratio at a significantly younger age than females, calculated by permuting matched age bins of the male and female groups to compute the null distribution of the difference in the peak ( $n(\text{permutations}) = 5000$ ,  $p < .001$ ). This p-value represents the proportion of permuted datasets that produced a mean difference between the peaks of the curves at least as extreme as the one observed from the actual data. These results were replicated using no smoothing (Suppl. Fig. S17), a smoothing kernel of 10 rather than 20 (Suppl. Fig. S18), as well as quantile bins with 20% (Suppl. Fig. S19–S20) instead of 10% of participants (range of median age at ratio maxima for females: 67.37–68.21 years; for males: 62.48–63.32 years).

In order to confirm the position of a change in slope for the



**Fig. 6.** Percent differences between left and right volumes. Bars represent the differences between left and right volumes, expressed as percentages of the average size of the regions, calculated by  $\frac{|\text{left volume} - \text{right volume}|}{\frac{\text{left volume} + \text{right volume}}{2}} * 100$ . Negative percentage differences correspond to larger right volumes, positive percent differences correspond to larger left volumes. Black bars indicate a significant corresponding *t*-test for absolute volume differences, at  $p < .001$  (Bonferroni-corrected  $\alpha = 0.05/7 = 0.007$ ).



**Fig. 7.** Percent differences between males and females. Bars represent the differences between volumes for men and women, expressed as percentages of the average size of the regions, calculated by  $\frac{|\text{male} - \text{female}|}{\frac{\text{volume male} + \text{volume female}}{2}} * 100$ . Negative percentage differences correspond to larger right volumes, positive percent differences correspond to larger left volumes. Black bars indicate a significant corresponding *t*-test for absolute volume differences at  $p < .001$  (Bonferroni-corrected  $\alpha = 0.05/8 = 0.006$ ).

trajectory of hippocampal volume, joinpoint regression (Kim et al., 2000) was applied. This method tests statistically for the presence of zero, or up to four inflection points (joinpoints) in the slope of mean hippocampal volume across age. The maximum number of joinpoints

tested for is selected based on the number of observations. There was one inflection point for both males and females (Fig. 9A–B). In females, there was a significant change in slope at age-group 64–65 years (change of slope  $\Delta m = -32.44 \text{ mm}^3$ ; age 95% CI [61–62 years, 67–68 years],  $p < .0001$ ). In males, the change in slope was significant at age-group 63–64 years (change of slope  $\Delta m = -26.45 \text{ mm}^3$  95% CI [58–59 years, 66–67 years],  $p < .0001$ ; see Table 3 for Model estimates).

For the rest of the grey matter, there was one joinpoint in the slope (total grey matter – bilateral hippocampus) across age for females (Fig. 9C), with a significant change in slope of  $\Delta m = -1273.98 \text{ mm}^3$  at age-group 56–57 years ( $t(21) = -2.98$ , 95% CI [53–54 years, 58–59 years],  $p = .007$ ). By contrast, there was also a small, but significant change in slope for the rest of grey matter across age in males at age-group 62–63 years ( $\Delta m = -754.14 \text{ mm}^3$ ,  $t(21) = -4.87$ , 95% CI [57–58 years, 66–67 years],  $p < .0001$ ) (Fig. 9D).

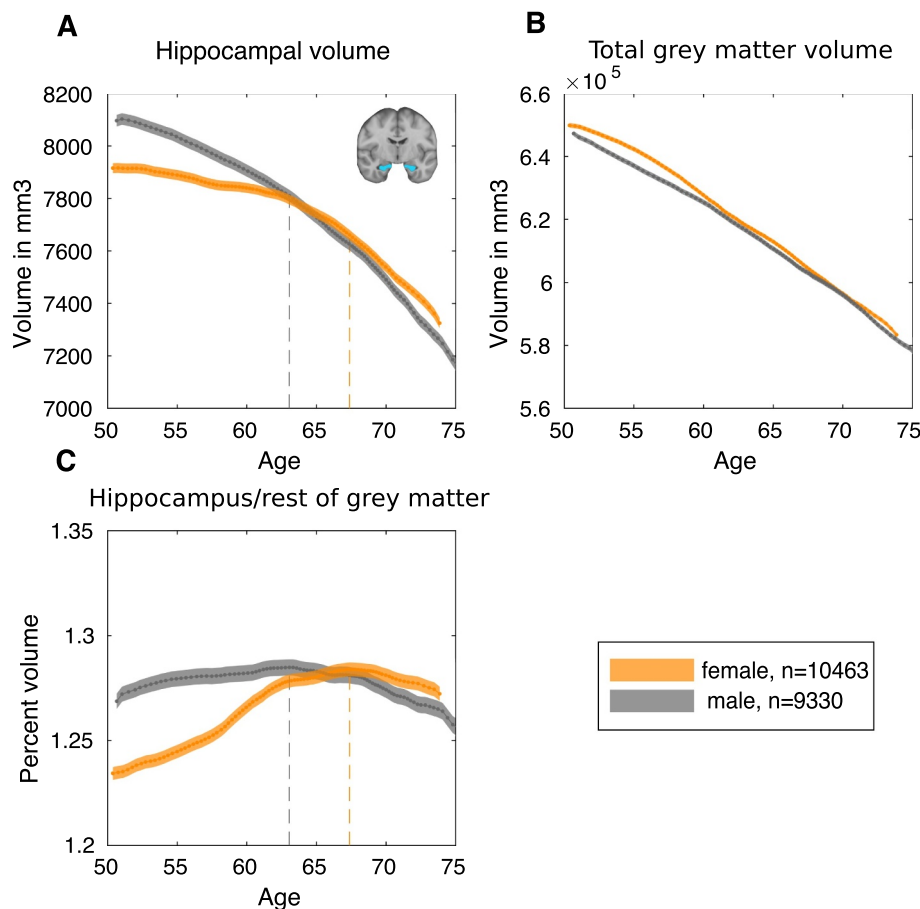
#### 3.4.4. Trajectories of other temporal lobe volumes with age

Next, we aimed to demonstrate that these inflections were specific to the hippocampus. Mean volumes (corrected for head size) were plotted for the remaining brain areas within the temporal lobe available from UK Biobank, namely parahippocampal gyrus, fusiform gyrus, superior temporal gyrus, middle temporal gyrus, inferior temporal gyrus, and temporal pole (Fig. 10). The same sliding-window and joinpoint regression methods were used. All these temporal lobe regions show a largely linear negative relationship with age. In keeping with this, and in contrast to the findings for the hippocampus, there were no significant joinpoints or changes in slope for either males or females in superior temporal gyrus, middle temporal gyrus, inferior temporal gyrus and fusiform gyrus ( $p > .05$ ). However, there was a significant joinpoint for both males and females in parahippocampal gyrus volume, similar to the joinpoints found for the hippocampus (females:  $\Delta m = -26.97 \text{ mm}^3$  at age-group 66–67 years,  $t(21) = -5.40$ , 95% CI [62–63 years, 70–71 years],  $p < .0001$ ; males:  $\Delta m = -30.81 \text{ mm}^3$  at age-group 59–60 years,  $t(21) = -3.77$ , 95% CI [56–57 years, 64–65 years],  $p = .001$ ) (Fig. 9D). There was also a significant joinpoint in volume of temporal pole in males, at age-group 55–56 years ( $\Delta m = -76.97 \text{ mm}^3$ ,  $t(21) = -2.57$ , 95% CI [53–54 years, 57–58 years],  $p = .02$ ).

In addition, the ages of the peak ratios with respect to the rest of grey matter differs between regions. For example, the ratios of superior (female:  $\text{bin}(\text{max ratio}) = 52.73 \text{ years} - 55.85 \text{ years}$ ; male:  $\text{bin}(\text{max ratio}) = 44.45 \text{ years} - 52.73 \text{ years}$ ) and middle temporal gyrus volume (female:  $\text{bin}(\text{max ratio}) = 49.30 \text{ years} - 53.11 \text{ years}$ ; male:  $\text{bin}(\text{max ratio}) = 48.30 \text{ years} - 53.11 \text{ years}$ ) to rest of grey matter peak in the youngest age groups of the sample. By contrast, ratios for inferior temporal gyrus (female:  $\text{bin}(\text{max ratio}) = 71.92 \text{ years} - 80.65 \text{ years}$ ; male:  $\text{bin}(\text{max ratio}) = 73.41 \text{ years} - 80.27 \text{ years}$ ), temporal pole (female:  $\text{bin}(\text{max ratio}) = 72.25 \text{ years} - 75.56 \text{ years}$ ) and parahippocampal volume (female:  $\text{bin}(\text{max ratio}) = 71.92 \text{ years} - 80.65 \text{ years}$ ; male:  $\text{bin}(\text{max ratio}) = 73.41 \text{ years} - 80.27 \text{ years}$ ) to the rest of the grey matter peak at the oldest age groups. The ratio of fusiform gyrus volume to the rest of grey matter volume does not show pronounced change over age. There were no significant differences in the age of peak ratio between males and females for any of these temporal lobe regions, calculated by permutation testing ( $n(\text{permutations}) = 5000$ , all  $p > .05$ ), except for fusiform gyrus ( $p = .004$ ).

#### 3.5. Effects of hypertension, smoking, higher BMI and lower education on brain volume

Before further analysing the demographic information, brain volumes were corrected for head size and scanning date as described above. As effects of hypertension status, Body-Mass-Index (BMI), education, and smoking status on brain volumes may also be confounded



**Fig. 8.** Trajectory of hippocampal volume with age. Dashed lines indicate points of maximum ratio. A. Mean bilateral hippocampal volume including standard errors as a function of age, corrected for head size. B. Mean total grey matter volume including standard errors as a function of age, corrected for head size. C. Mean hippocampal volume to rest of grey matter ratio including standard errors as a function of age.

by age, variance introduced by age in addition to head size and scanning date was regressed out for the following results on effects of hypertension, BMI and education only.

Participants with self-reported hypertension had significantly lower hippocampal ( $p < .001$ , Hedges'  $g = 0.06$ ) and total grey matter volumes ( $p < .001$ , Hedges'  $g = 0.17$ ), than participants without self-reported hypertension, though with small effect sizes (Suppl. Fig. S22). There was a significant, but very small negative correlation between BMI and hippocampal volume ( $r = -0.03$ ,  $p < .001$ ,  $n = 18,653$ ), as well as between BMI and total grey matter volume ( $r = -0.11$ ,  $p < .001$ ,  $n = 18,653$ ), with larger BMI being weakly associated with smaller volumes. There was also a small significant positive effect of level of education on average hippocampal volume ( $p < .001$ , Hedges'  $g = 0.07$ ,  $n = 19,646$ ). Higher education levels were associated with larger volumes than lower level education levels. There was no significant effect of education on total grey matter volume ( $p = .77$ ,  $n = 19,646$ , Suppl. Fig. S23).

Hippocampal volume was significantly smaller for current smokers than for people who have never smoked ( $p = .001$ , Hedges'  $g = 0.15$ ), and people who have previously smoked ( $p = .006$ , Hedges'  $g = 0.05$ ). Similarly, total grey matter volume was significantly reduced in current smokers compared to people who have never smoked ( $p < .0001$ , Hedges'  $g = 0.33$ ), and people who have previously smoked ( $p < .0001$ , Hedges'  $g = 0.17$ ). Current smoking was associated with lower volumes compared to previous smokers for total grey matter ( $p = .0001$ , Hedges'  $g = 0.15$ ), but not hippocampal volume ( $p > .008$ , Suppl. Fig. S24). P-values were corrected using the Bonferroni method with  $\alpha = 0.05/6$  multiple comparisons = 0.008. Results from the ANOVA also indicate a significant difference between hippocampal volumes of previous smokers and current smokers ( $F = 8.26$ ,  $p < .001$ ,  $n = 19,646$ ).

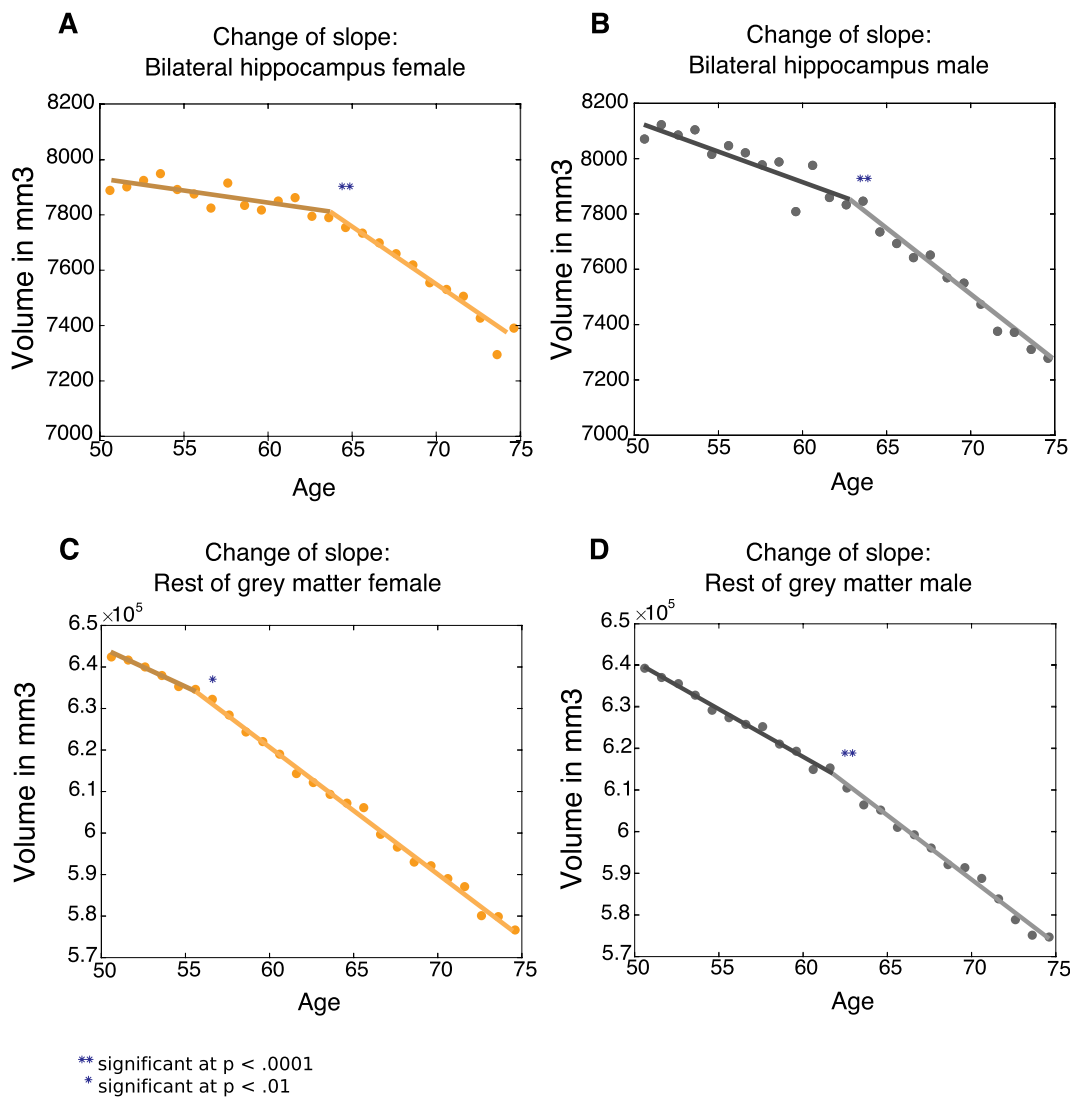
Supplementary Tables S1–3 present data for means, standard deviations, and significance tests of hypertension, BMI, education level, and smoking status for all temporal lobe volumes as well as total grey matter volume. Percent differences for all temporal lobe volumes are shown in Supplementary Figs. S22–24.

#### 4. Discussion

Data from the UK Biobank Imaging were analysed to provide normative information on hippocampal volume as a function of age using the largest sample size published to date ( $N = 19,793$ ). We used a model-free sliding window approach to examine absolute and relative hippocampal volume, as well as neighbouring temporal lobe volumes, as a function of age. This approach was deemed most appropriate as it makes few assumptions, and does not impose a linear relationship between brain volume and age. The analysis resulted in normative reference percentiles – or nomograms – for total grey matter and hippocampal volume along the age distribution, facilitating assessment of an individual's hippocampal volume in relation to their age-group (see our online tool at [www.win.ox.ac.uk/open-neuroimaging/analyse-your-data/hippocampal-volume-tool](http://www.win.ox.ac.uk/open-neuroimaging/analyse-your-data/hippocampal-volume-tool)).

For comparison, it is noteworthy that estimated hippocampal volumes from previous studies on patients with MCI or AD fall far below the 50th percentile in the healthy population presented here. For example, mean uncorrected hippocampal volume (bilateral) was  $5.55 \text{ cm}^3$  in an amnesic MCI group with an average age of 70.7 years (Vos et al., 2013). This would fall below the 2.5th percentile in our sample. In a sample of controls and patients with MCI and AD of similar mean age as in our sample (67 years, 71 years, and 67 years, respectively), controls would score between our 50th and 75th percentiles, while patients score below the 5th percentile (Henneman et al., 2009). Volumes





**Fig. 9.** Joinpoint analysis for hippocampal and total grey matter volume. A. Joinpoint (change of slope) in bilateral hippocampal volume over age in females B. Joinpoint bilateral hippocampus in males. C. Joinpoint in rest of grey matter volume (total grey matter – bilateral hippocampus) in females D. Joinpoint in rest of grey matter for males.

**Table 3**  
Model estimates from joinpoint regression of hippocampal volume.

	Female	Male
Degrees of freedom	21	21
Joinpoint	64–65 years	63–64 years
Joinpoint 95% Lower confidence level	61–62 years	58–59 years
Joinpoint 95% Upper confidence level	67–68 years	66–67 years
Slope change estimate	–32.44 mm <sup>3</sup>	–26.45 mm <sup>3</sup>
Slope change std error	4.04 mm <sup>3</sup>	4.94 mm <sup>3</sup>
Slope change test statistic <sup>a</sup>	–8.03	–5.35
Slope change p-value	< 0.0001	< 0.0001

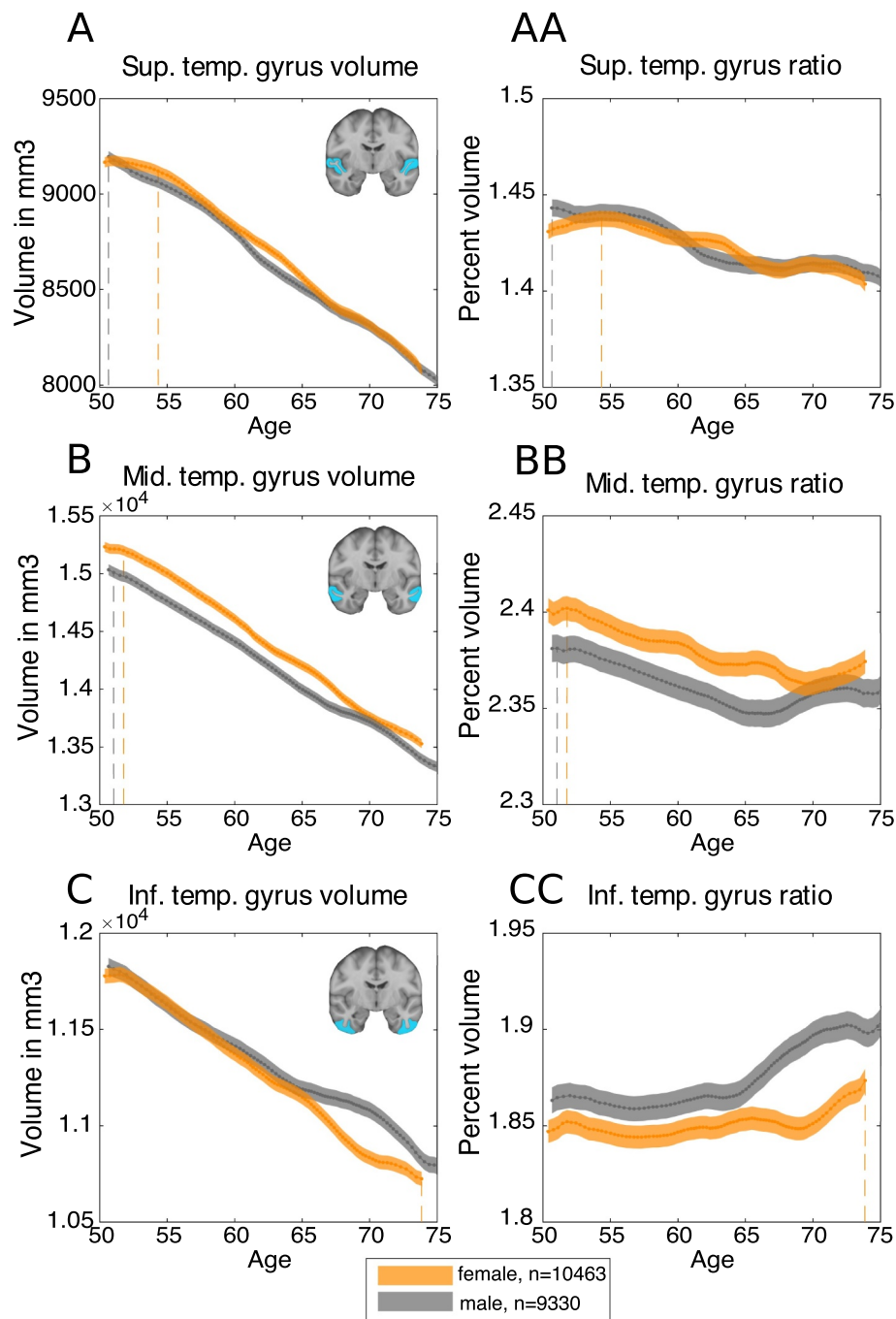
<sup>a</sup> The slope change test statistic is the parameter estimate divided by the standard error that follows a t distribution with 21 degrees of freedom.

compatible with our current results for healthy controls (3.57 cm<sup>3</sup>; mean age 75 years), but with far less dramatic reductions for MCI (3.26 cm<sup>3</sup>; mean age 73 years) and AD (2.97 cm<sup>3</sup>; mean age 72.6 years), were reported using data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (Mulder et al., 2014). These would likely correspond to healthy controls scoring around our 50th percentile, but with patients scoring between the 5th and 25th percentiles.

It is likely that different methods, including sample selection as well as image acquisition, image processing, and hippocampal segmentation tool used, underlie these discrepancies. For example, we found on average smaller estimates of hippocampal volume with FSL-FIRST than with FSL-FAST using atlas-based ROI selection. Another study also reported significant differences between FSL-FIRST, FreeSurfer, and manual segmentation (Mulder et al., 2014). However, it is less clear how more severe deviations from acquisition and analysis protocols may affect volume estimations. These issues highlight the importance of large datasets and unified analysis pipelines. Thus, before efforts to harmonize analysis across MRI datasets solve some of these concerns, the nomograms presented here provide a robust comparator, when used with estimates obtained using the same, or similar, acquisition protocols, hardware and pre-processing pipeline.

It can also currently not easily be disambiguated whether a low percentile on the nomogram is due to pathological atrophy or is simply constitutional, e.g., due to an unusually small volume from birth. However, future research with longitudinal data from the UK Biobank may facilitate the resolution of some of these constraints.

Associations of hippocampal volume with education, smoking, and hypertension were significant, but effect sizes were small, and deconfounding for these variables did not effectively change any of the



**Fig. 10.** Trajectory of temporal lobe volumes with age, corrected for head-size.

Dashed lines indicate points of maximum ratio. A. Bilateral superior temporal gyrus volume. AA. Bilateral superior temporal gyrus ratio to rest of grey matter. B. Bilateral middle temporal gyrus volume. BB. Bilateral middle temporal gyrus ratio to rest of grey matter. C. Bilateral inferior temporal gyrus volume. CC. Bilateral inferior temporal gyrus ratio to rest of grey matter. D. Bilateral fusiform gyrus volume. DD. Bilateral fusiform gyrus ratio to rest of grey matter. E. Bilateral parahippocampal gyrus volume. EE. Bilateral parahippocampal gyrus ratio to rest of grey matter. F. Bilateral temporal pole volume. FF. Bilateral temporal pole ratio to rest of grey matter.

reported results. Since significant findings with small effect sizes are not necessarily practically relevant with big data, we decided not to correct for these factors in the calculations of percentiles in favour of generalisability. However, it should be noted that participants in UK Biobank appear to smoke less, have lower BMIs, and are generally healthier than would be expected in the general population of the United Kingdom (Fry et al., 2017).

The results of two different types of analysis revealed a slight acceleration of hippocampal volume loss around age 60–65 years for

females, whereas for males the rate of hippocampal volume loss may increase earlier around 50 years (Figs. 7 and 8). In addition, for both women and men, there was an increase in rate of hippocampal volume loss relative to the rest of the grey matter from around ages 67 and 63 years, respectively. Thus, hippocampal volume declined slower than the rest of grey matter until around age 67 in women, and age 63 in men, after which hippocampal volume declined faster than the rest of grey matter. This may indicate a particular vulnerability of the hippocampus in ageing, as this effect was specific to the hippocampus and

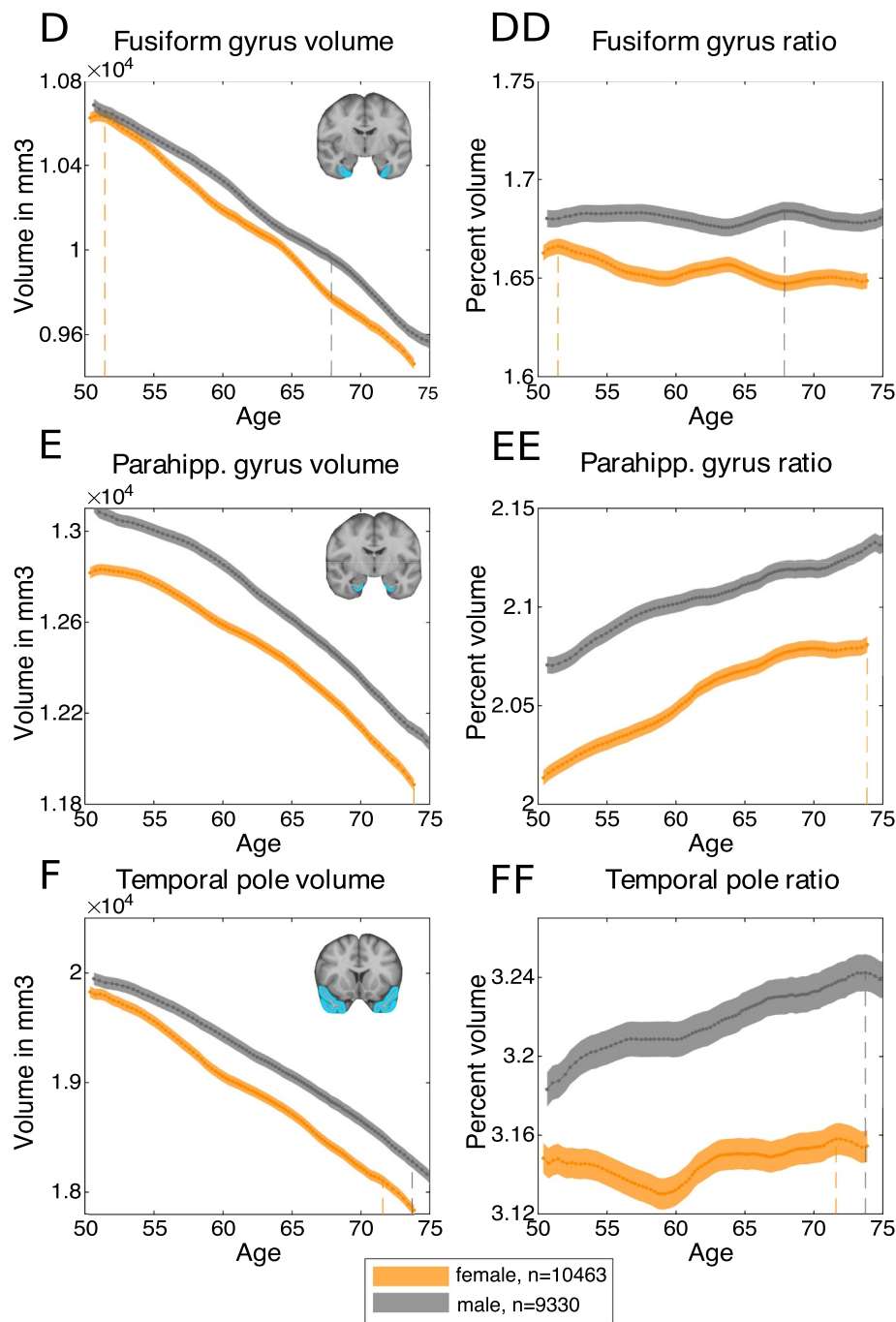


Fig. 10. (continued)

was not found to this extent in neighbouring brain areas such as parahippocampal gyrus or temporal gyrus (Fig. 9). While our method does not allow for precise age estimations, we show a robust age-range for the ratio peak with different window bins and smoothing kernels.

A previous cross-sectional study with results from 1100 participants reported a similar acceleration of rate of hippocampal volume loss (Fjell et al., 2013). Relationships of age and several brain volumes were analysed using a nonparametric smoothing spline approach. The authors reported an acceleration of rate of hippocampal volume loss starting around 50 years of age, with a marked increase in rate of hippocampal volume loss around 60 years. In agreement with our findings, no such acceleration point was found in grey matter. However, in this study no sex differences in these curves were calculated. In contrast, other cross-sectional studies using linear regression to

estimate associations of age with hippocampal volume found an acceleration of volume loss at a slightly later age of 72 years (Zhang et al., 2010), or no acceleration point (Knoops et al., 2012).

Based on the trajectory of hippocampal volume across age presented here however, we argue that a linear fit is not the most appropriate approach to measure this association. Importantly, the trajectories of other temporal lobe volumes, as well as the rest of grey matter across age do not show the same acceleration of volume loss as observed in hippocampal volume. This may underlie the specific vulnerability of hippocampal volume loss in older age. Yet, the findings presented here vary slightly depending on method used (e.g. sliding window or joinpoint), and need to be validated with longitudinal rather than cross-sectional data. For example, results of the joinpoint regression indicate that if there were a change in slope for hippocampal volume across age,

then it would occur at 64–65 years, and would be significant. However, joinpoint regression cannot be applied to establish whether a change in slope is present or absent, as the method tends to identify a join point even if there is a continuous transition. In addition, the current data do not allow for conclusions of whether those with the fastest decline, or those with the lowest baseline volume, are at more risk of developing dementia. With the availability of longitudinal health outcomes, future studies may also explore the role of the dementia-associated APOE  $\epsilon 4$  allele on hippocampal volume across age.

In the absence of differential age-effects on grey matter volume between males and females, we found significantly larger mean total grey matter, and middle temporal gyrus volumes in females after correcting for head size and age. However, effect sizes were small. On the other hand, males had moderately larger parahippocampal and temporal pole volumes. The data presented here also indicate hemispheric asymmetry of the tested brain volumes. The right hippocampus was slightly, but significantly, larger than the left hippocampus for both men and women. Right-larger-than-left asymmetry was also found in the superior and inferior temporal gyri. In contrast, there was considerable left-larger-than-right asymmetry for both the parahippocampal and fusiform gyrus. The observed asymmetries, and the differences in asymmetry direction, may be the result of noise in the MRI signals and imperfections of the automated volume segmentation and quality control tools used here. For example, the algorithm that the segmentation uses to determine which voxel is classified as 'parahippocampal' versus 'hippocampal' may be the reason for the differences in asymmetry direction between the two volumes. Thus, while these results suggest some structural differences between hemispheres, such as asymmetry of convolutions, it is unclear what these differences entail precisely. However, the right-larger-than-left asymmetry of the hippocampus has been reported previously in smaller cohorts of younger and older adults (Zhang et al., 2010; Wellington et al., 2013; Pedraza et al., 2018), as well as in patients with AD and MCI (Shi et al., 2009). A recent study reporting on 400 participants from the ADNI dataset also found increasing right-larger-than-left hippocampus asymmetry going from healthy controls, to patients with MCI, to patients with AD (Alberich-Bayarri et al., 2018). As the left hippocampus is smaller even in healthy ageing, the increasing asymmetry in the course of AD may be explained by greater vulnerability to pathology of the left compared to the right hippocampus. What is less clear is how the left-larger-than-right asymmetries of the parahippocampal and fusiform gyri may be affected by ageing and AD pathology.

Functional lateralisation has been shown for a number of brain processes, such as language (Vigneau et al., 2006), face perception (Zhen et al., 2015), or visual processing (Zhen et al., 2017), which may explain the asymmetries observed here in other areas such as the inferior and superior temporal gyri. However, the precise nature of the relationship between brain structural lateralisation and function remains unclear (Bishop, 2013; Batista-García-Ramó and Fernández-Verdecia, 2018), highlighting the importance of considering confounding factors for the analysis of norm values. We therefore provide separate nomograms for left and right hippocampus, as well as for males and females.

## 5. Conclusion

Analysis of 19,793 generally healthy participants in the UK Biobank revealed effects of age, sex, and hemisphere on selected temporal lobe volumes. The data provide normative values for hippocampal and total grey matter volume as a function of age for reference in research and clinical settings, based on an unprecedented sample size. These norm values may be used together with automated percentile estimation tools such as the one described here to provide a rapid, but objective, evaluation of the patient's hippocampal volume status. While the current findings are based on cross-sectional data, the longitudinal health outcomes that will become available in UK Biobank over the next years

will add invaluable information to the role of hippocampal atrophy in ageing.

## Acknowledgements

### Funding

This research was conducted using the UK Biobank resource (application ID: 32011). UK Biobank was established by the Wellcome Trust, the Medical Research Council, the United Kingdom Department of Health, and the Scottish Government. The UK Biobank has also received funding from the Welsh Assembly Government, the British Heart Foundation, and Diabetes UK.

This work was supported by the Wellcome Trust (206330/Z/17/Z to M.H.), the Medical Research Council (DPhil Studentship to L.N. and a Clinician Scientist Fellowship to SGM), SMS, FA-A, MJ, and CEM acknowledge support from the National Institute for Health Research (NIHR) Oxford Health Biomedical Research Centre (BRC), a partnership between Oxford Health NHS Foundation Trust and the University of Oxford. The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nicl.2019.101904>.

## References

- Adriano, F., Caltagirone, C., Spalletta, G., 2012. Hippocampal volume reduction in first-episode and chronic schizophrenia: A review and meta-analysis. *Neurosci.* 18, 180–200.
- Ahmed, R.M., Paterson, R.W., Warren, J.D., 2014. Biomarkers in dementia: clinical utility and new directions. *J. Neurol. Neurosurg. Psychiatry* 85, 1426–1434.
- Alberich-Bayarri, Á., Junque, C., Zhou, J., Sarica, A., Vasta, R., Novellino, F., Vaccaro, M.G., Ceresa, A., Quattrone, A., 2018. MRI asymmetry index of hippocampal subfields increases through the continuum from the mild cognitive impairment to the Alzheimer's disease. *Front. Neurosci.* 1, 576.
- Alfaro-Almagro, F., Jenkinson, M., Bangerter, N.K., Andersson, J.L.R., Griffanti, L., Douaud, G.E., Sotiropoulos, S.N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., Vidaurre, D., Webster, M., McCarthy, P., Rorden, C., Daducci, A., Alexander, D.C., Zhang, H., Dragonu, I., Matthews, P.M., Miller, K.L., Smith, S.M., 2018. Image processing and quality control for the first 10,000 brain imaging datasets from UK biobank. *Neuroimage* 166, 400–424.
- Barnes, J., Bartlett, J.W., van de Pol, L.A., Loy, C.T., Scallan, R.I., Frost, C., Thompson, P., Fox, N.C., 2009. A meta-analysis of hippocampal atrophy rates in Alzheimer's disease. *Neurobiol. Aging* 30, 1711–1723.
- Barnes, J., Ridgway, G.R., Bartlett, J., Henley, S.M.D., Lehmann, M., Hobbs, N., Clarkson, M.J., Macmanus, D.G., Ourselin, S., Fox, N.C., 2010. Head size, age and gender adjustment in MRI studies: a necessary nuisance? *Neuroimage* 53, 1244–1255.
- Batista-García-Ramó, K., Fernández-Verdecia, C.I., 2018. What we know about the brain structure-function relationship. *Behav. Sci.* 8.
- Bishop, D.V.M., 2013. Cerebral asymmetry and language development: cause, correlate, or consequence? *Science* 340, 1230531.
- Button, K.S., A Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., J Robinson, E.S., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376.
- Campbell, S., Marriott, M., Nahmias, C., Macqueen, G.M., 2004. Reviews and overviews lower hippocampal volume in patients suffering from depression: A meta-analysis. *Am. J. Psychiatry* 161, 598–607.
- Chakos, M.H., Schobel, S.A., Gu, H., Gerig, G., Bradford, D., Charles, C., Lieberman, J.A., 2005. Duration of illness and treatment effects on hippocampal volume in male patients with schizophrenia. *Br. J. Psychiatry* 186, 26–31.
- Chételat, G., Landeau, B., Eustache, F., Mézenge, F., Viader, F., De Lea Sayette, V., Desgranges, B., Baron, J.-C., 2005. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: A longitudinal MRI study. *Neuroimage* 27, 934–946.
- Choe, Y.M., Kim, K.W., Jhoo, J.H., Ryu, S.H., Seo, E.H., Sohn, B.K., Byun, M.S., Bak, J.H., Lee, J.M., Yun, H.J., Han, M.I., Woo, J.I., Lee, D.Y., 2016. Multicenter, randomized, placebo-controlled, double-blind clinical trial of escitalopram on the progression-delaying effects in Alzheimer's disease. *Int. J. Geriatr. Psychiatry* 31, 731–739.
- Colle, R., Dupong, I., Colliot, O., Deflesselle, E., Hardy, P., Falissard, B., Ducreux, D., Chupin, M., Corbule, E., Ne Dupong, I., 2016. Smaller hippocampal volumes predict lower antidepressant response/remission rates in depressed patients: A meta-analysis. *World J. Biol. Psychiatry* 19, 360–367.
- Finke, C., Prüss, H., Heine, J., Reuter, S., Kopp, U.A., Wegner, F., Then Bergh, F., Koch, S., Jansen, O., Münte, T., Deuschl, G., Rupprecht, K., Stöcker, W., Wandinger, K.-P., Paul,



- F., Bartsch, T., 2017. Evaluation of cognitive deficits and structural hippocampal damage in encephalitis with leucine-rich, glioma-inactivated 1 antibodies. *JAMA Neurol.* 74, 50.
- Fjell, A.M., Westlye, L.T., Grydeland, H., Amlie, I., Espeseth, T., Reinvang, I., Raz, N., Holland, D., Dale, A.M., Walhovd, K.B., 2013. Critical ages in the life course of the adult brain: nonlinear subcortical aging. *Neurobiol. Aging* 34, 2239–2247.
- Frankó, E., Joly, O., Herholz, K., 2013. Evaluating Alzheimer's disease progression using rate of regional hippocampal atrophy for the Alzheimer's disease neuroimaging initiative. *PLoS ONE* 8.
- Fraser, M.A., Shaw, M.E., Cherbuin, N., 2015. A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing. *Neuroimage* 112, 364–374.
- Frisoni, G.B., Fox, N.C., Jack, C.R., Scheltens, P., Thompson, P.M., 2010. The clinical use of structural MRI in Alzheimer disease. *Nat. Rev. Neurol.* 6, 67–77.
- Fry, A., Littlejohns, T.J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., Allen, N.E., 2017. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. *Am. J. Epidemiol.* 186, 1026–1034.
- Henneman, W.J.P., Sluimer, J.D., Barnes, J., Van Der Flier, W.M., Sluimer, I.C., Fox, N.C., Scheltens, P., Vrenken, H., Barkhof, F., 2009. Hippocampal atrophy rates in Alzheimer disease added value over whole brain volume measures. *Neurology* 72, 999–1007.
- Hentschke, H., Stuetgen, M.C., 2011. Computation of measures of effect size for neuroscience data sets. *Eur. J. Neurosci.* 34, 1887–1894.
- Ioannidis, J.P.A., 2011. Excess significance bias in the literature on brain volume abnormalities. *Arch. Gen. Psychiatry* 68, 773–780.
- Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841.
- Kandiah, N., Zainal, N.H., Narasimhalu, K., Chander, R.J., Ng, A., Mak, E., Au, W.L., Sitoh, Y.Y., Nadkarni, N., Tan, L.C.S., 2014. Hippocampal volume and white matter disease in the prediction of dementia in Parkinson's disease. *Parkinsonism Relat. Disord.* 20, 1203–1208.
- Kim, H., Fay, M., Feuer, E., Midthune, D., 2000. Permutation tests for joinpoint regression with applications to cancer rates. *Stat. Med.* 19, 335–351 (correction:2001;20:655).
- Kishi, T., Matsunaga, S., Oya, K., Ikuta, T., Iwata, N., 2015. Protection against brain atrophy by anti-dementia medication in mild cognitive impairment and Alzheimer's disease: meta-analysis of longitudinal randomized placebo-controlled trials. *Int. J. Neuropsychopharmacol.* 1–7.
- Knoops, A.J.G., Gerritsen, L., van der Graaf, Y., Mali, W.P.T.M., Geerlings, M.I., 2012. Loss of entorhinal cortex and hippocampal volumes compared to whole brain volume in normal aging: the SMART-Medea study. *Psychiatry Res. Neuroimaging* 203, 31–37.
- Koenig, K.A., Sakaie, K.E., Lowe, M.J., Lin, J., Stone, L., Bermel, R.A., Beall, E.B., Rao, S.M., Trapp, B.D., Phillips, M.D., 2014. Hippocampal volume is related to cognitive decline and fornical diffusion measures in multiple sclerosis. *Magn. Reson. Imaging* 32, 354–358.
- Kril, J.J., Patel, S., Harding, A.J., Halliday, G.M., 2002. Neuron loss from the hippocampus of Alzheimer's disease exceeds extracellular neurofibrillary tangle formation. *Acta Neuropathol.* 103, 370–376.
- Manohar, S.G., 2019. Matplotlib: MATLAB Tools for Plotting, Data Analysis, Eye Tracking and Experiment Design (Public).
- MATLAB R2017b, The MathWorks, Inc., Natick, Massachusetts, United States
- Mielke, M.M., Okonkwo, O.C., Oishi, K., Mori, S., Tighe, S., Miller, M.I., Brown, T., Albert, M., Lyketsos, C.G., 2012. Fornix integrity and hippocampal volume predict memory decline and progression to Alzheimer's disease. *Alzheimers Dement.* 105–113.
- Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L.R., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK biobank prospective epidemiological study. *Nat. Neurosci.* 19, 1523–1536.
- Morimoto, T., Matsuda, Y., Matsuoka, K., Yasuno, F., Ikebuchi, E., Kameda, H., Taoka, T., Miyasaka, T., Kichikawa, K., Kishimoto, T., 2018. Computer-assisted cognitive remediation therapy increases hippocampal volume in patients with schizophrenia: a randomized controlled trial. *BMC Psychiatry* 18.
- Mulder, E.R., de Jong, R.A., Knol, D.L., van Schijndel, R.A., Cover, K.S., Visser, P.J., Barkhof, F., Vrenken, H., 2014. Hippocampal volume change measurement: quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST. *Neuroimage* 92, 169–181.
- Nord, C.L., Valton, V., Wood, J., Roiser, J.P., 2017. Power-up: A reanalysis of “power failure” in neuroscience using mixture Modeling. *J. Neurosci.* 37, 8051–8061.
- Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M., 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 56, 907–922.
- Pedraza, O., Bowers, R., Gilmore, R., 2018. Asymmetry of the hippocampus and amygdala in MRI volumetric measurements of normal adults. *J. Int. Neuropsychol. Soc.* 10, 664–678.
- Reyes, A., Holden, H.M., Chang, Y.-H.A., Uttarwar, V.S., Sheppard, D.P., Deford, N.E., DeJesus, S.Y., Kansal, L., Gilbert, P.E., McDonald, C.R., 2018. Impaired spatial pattern separation performance in temporal lobe epilepsy is associated with visuospatial memory deficits and hippocampal volume loss. *Neuropsychologia* 111, 209–215.
- Schuff, N., Woerner, N., Boreta, L., Kornfield, T., Shaw, L.M., Trojanowski, J.Q., Thompson, P.M., Jack, C.R., Weiner, M.W., 2009. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain* 132, 1067–1077.
- Shi, F., Liu, B., Zhou, Y., Yu, C., Jiang, T., 2009. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies. *Hippocampus* 19, 1055–1064.
- Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, A., De Stefano, N., 2002. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* 17, 479–489.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niaz, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219.
- Stasinopoulos, D.M., Rigby, R.A., 2007. Generalized additive models for location scale and shape (GAMLSS) in R. *J. Stat. Softw.* 23.
- Takao, H., Hayashi, N., Ohtomo, K., 2011. Effect of scanner in longitudinal studies of brain volume changes. *J. Magn. Reson. Imaging* 34, 438–444.
- Videbech, P., Ravnkilde, B., 2004. Hippocampal volume and depression: A meta-analysis of MRI studies. *Am. J. Psychiatry* 161, 1957–1966.
- Vigneau, M., Beaucoeur, V., Hervé, P.Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., Tzourio-Mazoyer, N., 2006. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30, 1414–1432.
- Vos, S.J.B., Van Rossum, I.A., Verhey, F., Knol, D.L., Soininen, H., Wahlund, L.O., Hampel, H., Tsolaki, M., Minthon, L., Frisoni, G.B., Froelich, L., Nobili, F., Van Der Flier, W., Blennow, K., Wolz, R., Scheltens, P., Visser, P.J., 2013. Prediction of Alzheimer disease in subjects with amnesic and nonamnesic MCI. *Neurology* 80, 1124–1132.
- Wellington, R.L., Bilder, R.M., Napolitano, B., Szeszko, P.R., 2013. Effects of age on prefrontal subregions and hippocampal volumes in young and middle-aged healthy humans. *Hum. Brain Mapp.* 34, 2129–2140.
- World Health Organization Multicentre Growth Reference Study Group, 2006. WHO child growth standards based on length/height, weight and age. *Acta Paediatr. Suppl.* 450, 76–85.
- Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45–57.
- Zhang, Y., Qiu, C., Lindberg, O., Bronge, L., Aspelin, P., Bäckman, L., Fratiglioni, L., Wahlund, L.-O., 2010. Acceleration of hippocampal atrophy in a non-demented elderly population: the SNAC-K study. *Int. Psychogeriatr.* 22, 14–25.
- Zhen, Z., Yang, Z., Huang, L., Kong, X.-Z., Wang, X., Dang, X., Huang, Y., Song, Y., Liu, J., 2015. Quantifying interindividual variability and asymmetry of face-selective regions: A probabilistic functional atlas. *Neuroimage* 113, 13–25.
- Zhen, Z., Kong, X.-Z., Huang, L., Yang, Z., Wang, X., Hao, X., Huang, T., Song, Y., Liu, J., 2017. Quantifying the variability of scene-selective regions: Interindividual, inter-hemispheric, and sex differences. *Hum. Brain Mapp.* 38, 2260–2275.