This research was funded in whole, or in part, by the Wellcome Trust 109067/Z/15/AI . For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission

Title: T2 heterogeneity as an *in vivo* marker of microstructural integrity in medial temporal lobe subfields in ageing and mild cognitive impairment

Alfie R. Wearn* [a], Volkan Nurdal [a], Esther Saunders-Jennings [a], Michael J. Knight [b], Christopher R. Madan [d], Sean-James Fallon [e], Hanna K. Isotalus [a], Risto A. Kauppinen [f], Elizabeth J. Coulthard [a c]

Affiliations

a Bristol Medical School, University of Bristol, Bristol, UK

b School of Psychological Science, University of Bristol, Bristol, UK

c Clinical Neurosciences, North Bristol NHS Trust, Bristol, UK

d School of Psychology, University of Nottingham, Nottingham, UK

e National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and University of Bristol, Bristol, UK

f Faculty of Engineering, University of Bristol, Bristol, UK

*Corresponding author: Alfie R. Wearn

Corresponding author's address: Institute of Clinical Neurosciences, Learning & Research Bldg, Southmead Hospital, BS10 5NB

Corresponding author's e-mail address: Alfie.wearn@bristol.ac.uk

Corresponding author's phone number: +44 (0)7908038054

Keywords

Magnetic resonance imaging, Alzheimer's disease, Early diagnosis, Ageing, Hippocampal subfields, Medial Temporal Lobe, T2 relaxometry, T2 heterogeneity

ABSTRACT

A better understanding of early brain changes that precede loss of independence in diseases like Alzheimer's disease (AD) is critical for development of disease-modifying therapies. Quantitative MRI, such as T2 relaxometry, can identify microstructural changes relevant to early stages of pathology. Recent evidence suggests heterogeneity of T2 may be a more informative measure of early pathology than absolute T2. Here we test whether T2 markers of brain integrity precede the volume changes we know are present in established AD and whether such changes are most marked in medial temporal lobe (MTL) subfields known to be most affected early in AD. We show that T2 heterogeneity was greater in people with mild cognitive impairment (MCI; n=49) compared to healthy older controls (n=99) in all MTL subfields, but this increase was greatest in MTL cortices, and smallest in dentate gyrus. This reflects the spatio-temporal progression of neurodegeneration in AD. T2 heterogeneity in the entorhinal cortex also predicted cognitive decline over a year in people with MCI, where measures of volume or T2 in any other subfield or whole hippocampus could not. Increases in T2 heterogeneity in MTL cortices may reflect localised pathological change and may present as one of the earliest detectible brain changes prior to atrophy. Finally, we describe a mechanism by which memory, as measured by accuracy and reaction time on a paired associate learning task, deteriorates with age. Age-related memory deficits were explained in part by lower subfield volumes, which in turn were directly associated with greater T2 heterogeneity. We propose that tissue with high T2 heterogeneity represents extant tissue at risk of permanent damage but with the potential for therapeutic rescue. This has implications for early detection of neurodegenerative diseases and the study of brain-behaviour relationships.

INTRODUCTION

Accurate early diagnosis of Alzheimer's disease (AD) is likely a necessity for development of disease-modifying therapies (Cummings et al., 2014, Alzheimer's Association, 2015). Manifestation of cognitive symptoms, although required for clinical diagnosis, is a relatively late stage in the pathological process (Jack et al., 2010). Thus, clinical interventions after the appearance of cognitive deficits may be too late to restore brain health. A better understanding of the brain changes that precede loss of daily independence will help design early markers.

Structural and quantitative MRI show promise in their ability to identify changes in the brain that indicate early Alzheimer's pathology. Identifying which people with mild cognitive impairment (MCI) will progress to AD dementia has been shown to be possible by measuring the volume of subfields within the medial temporal lobe (MTL) (Chételat et al., 2008, de Flores et al., 2015a, de Flores et al., 2015b, deToledo-Morrell et al., 2004, Apostolova et al., 2006, Apostolova et al., 2010). In these groups, large changes in volume tend to indicate significant, and likely irreversible, atrophy. Smaller scale microstructural changes that occur prior to volume loss could help to identify patients in which such a treatment is to be optimally effective.

Recently, we demonstrated that the distribution width of T2 relaxation time (T2 heterogeneity) in the hippocampus predicted cognitive decline over a year in a group of people with MCI (Wearn et al., 2020a). We propose that T2 distribution widens because of different pathological hallmarks having opposing effects on T2, causing increased apparent heterogeneity without any change in absolute T2 (distribution midpoint). For example, non-haem iron, oligomers and plaques of β-amyloid (Aβ), and neurofibrillary tangles (NFTs) which build up around the MTL in early Alzheimer's disease (Braak and Braak, 1991, Braak and Braak, 1995, Selkoe and Hardy, 2016, Smith et al., 2010) all cause T2 to decrease (Meadowcroft et al., 2015, House et al., 2008). In contrast, tissue alterations preceding necrosis cause cell membrane breakdown and oedema which increase the motility of water within a

given region, subsequently causing T2 to increase (Laakso et al., 1996, Symms et al., 2004). Gliosis, another hallmark of brain injury that commonly follows neurodegeneration is also indicated by increased T2 relaxation time (Briellmann et al., 2002, Ingelsson et al., 2004, Lee et al., 2013). These opposing factors necessitate examination of the heterogeneity of T2, rather than midpoint, for more accurate identification of microstructural impairment. This, we propose, is a reason for the lack of clear consensus from previous studies of T2 in AD, which exclusively look at absolute T2 (see Tang et al. (2018) for a review).

Our previous research focused on T2 changes in the hippocampus as a whole (Wearn et al., 2020a). However, the hippocampus is not a uniform structure, rather, it comprises cytoarchitectonic subfields with distinct cellular structure, connectivity, functionality, and disease susceptibility (Duvernoy et al., 2013). NFTs first build up in the transentorhinal region of the MTL (Braak and Braak, 1995), which roughly corresponds to Brodmann area 35 (BA35), but also includes some of the lateral portion of entorhinal cortex (EC) (Xie et al., 2018). They then spread through EC, then to CA1, subiculum, other CA regions and finally dentate gyrus (DG) (Braak and Braak, 1995, Braak and Braak, 1991, Fukutani et al., 1995, Fukutani et al., 2000). Many of these changes occur even before symptom onset (Braak and Braak, 1991, Jack et al., 2003, Fukutani et al., 1995), highlighting their potential as prodromal markers. Reflective of histopathology, hippocampal volume loss due to AD is widespread across the hippocampus but is generally non-uniform across subfields, with most atrophy seen in CA1 (Adler et al., 2018, La Joie et al., 2013, Mueller et al., 2010, Wolk et al., 2017, Kerchner et al., 2012, Sarazin et al., 2010, Frisoni et al., 2008, Frisoni et al., 2006). Volume loss is less severe in the hippocampus of people with MCI, often restricted to CA1 and, in many cases, the subiculum (Pluta et al., 2012, Mueller et al., 2010, Tang et al., 2014), with some evidence for a relative sparing of the stratum radiatum/lacunosum/moleculare (Su et al., 2018). This atrophy pattern has even been shown in people who subjectively report cognitive decline but who have normal cognition as measured by standard cognitive tests (Perrotin et al., 2015). For a review see de Flores et al. (2015b).

Literature on quantitative T2 in MTL subfields in the context of AD is limited to very few *ex vivo* studies (Huesgen et al., 1993, Antharam et al., 2012). As with the rest of the literature, these papers are focused on absolute T2. Antharam et al. (2012) do note that the distribution width of T2 within the main hippocampal subfields (CA4-DG and CA1-3) is wider in slices from AD patients than age-matched controls. However, they do not perform detailed analyses of the differences of distribution width between subfields. In a publication of pilot data from our group, Knight et al. (2019) concluded that T2 heterogeneity in MTL subfields can improve accuracy in distinguishing between healthy controls, those with MCI and Alzheimer's disease patients. We know of no other studies that have explored quantitative T2 in subfields of the MTL *in vivo*.

The analyses in this paper are presented in two parts. In the first part, we test whether differences in T2 heterogeneity between subfields could distinguish healthy aging from MCI – important when considering T2 as a clinic tool to guide prognosis in MCI. We also analyse absolute T2 (distribution midpoint) to verify that it is heterogeneity and not absolute values that shift. Hypotheses:

1. The effect size of T2 heterogeneity increase in MCI (compared to healthy controls) will differ by subfield, reflective of the spatio-temporal progression of neurodegeneration in AD. Accordingly, we expect to see greatest T2 heterogeneity in MTL cortical regions (BA35 and EC), followed by CA and SUB regions, and the least amount of heterogeneity in DG.

2. T2 heterogeneity in MTL subfields will better predict cognitive decline than in whole hippocampus in people with mild cognitive impairment.

In the second part, we use path analysis to demonstrate the likely temporal sequence of neuroanatomical and behavioural changes in ageing. These are important to understand so that we track the right process at the right disease stage. Hypotheses:

3. Greater T2 heterogeneity indicates early damage that will lead to macroscopic structural change, and therefore will statistically mediate the relationship between age and volume of MTL subfields.

4. Subfield volume is indicative of macroscopic structural change, and therefore will mediate the relationship between T2 heterogeneity and cognitive performance.

MATERIAL AND METHODS

The following methods are adapted from those presented by Wearn et al. (2020a).

The analyses in this paper combine data from two prospective longitudinal studies similar in cohort demographics and study design. No participants took part in both studies. Both studies are detailed in the following section. Where data collected are not identical between cohorts, we have normalised equivalent metrics within cohort and combined data after normalisation.

Participants

Participants fulfilling the Petersen criteria (Albert et al., 2011) for diagnosis of MCI were recruited to both studies (Study 1: n=30, Study 2: n=29). Healthy older people (HC), with no history of memory problems or significant neurological disorders were recruited as controls to each study (Study 1: n=61; Study 2: n=56). All healthy controls had Montreal Cognitive Assessment (MoCA) $\geq$ 26 (study 1) or Addenbrookes Cognitive Examination 3 (ACE-III) $\geq$ 88 (study 2). 7 participants originally recruited as healthy controls in study 1 were found to have MoCA scores of < 26, so were reclassified as MCI (given the high sensitivity and specificity of the MoCA for detecting MCI at this threshold; 90% and 100%, respectively (Nasreddine et al., 2005)).

Subjects for both studies were recruited from local GP surgeries and memory clinics in the Bristol area (having received MCI diagnoses or reported memory problems), Join Dementia Research, Avon and Wiltshire Mental Health Partnership's Everyone Included system, an in-house database of volunteers, replies to poster adverts or through word of mouth. All patients provided informed written consent prior to testing as according to the Declaration of Helsinki. Ethical approval was given by Frenchay NHS Research Ethics Committee.

The current analyses included all participants who had both volumetry and T2 relaxometry data for hippocampal subfields, study 1 n=91 (50 HC, 30 MCI), study 2 n=66 (49 HC, 19 MCI). See tables 1

and 2 for demographic details. A total of 20 MCI participants were followed-up after one-year (10 from each study), representing a relatively high dropout rate from the MCI cohort.

Cognitive testing

Cognitive function was tested at baseline and follow-up using the MoCA in study 1 and the ACE-III in study 2.

Participants in both studies carried out the paired associates learning (PAL) task of the CANTAB toolbox which has shown high sensitivity to cognitive impairment and daily functioning in dementia (Égerházi et al., 2007).

Imaging parameters

Scans for both studies were acquired on a Siemens Magnetom Skyra 3T system equipped with a parallel transmit body coil and a 32-channel head receiver array coil. The two studies used similar, but slightly different scanning protocols.

Study 1

This protocol has been previously described by Knight et al. (2019). The imaging protocol included a 3D T1-weighted whole-brain magnetization prepared rapid acquisition gradient-echo (MPRAGE) and 2D multi-contrast multi-spin-echo (CPMG).

MPRAGE: Coronal, whole-brain, repetition time (TR) 2200 ms, Echo Time (TE) 2.42 ms, Inversion time (TI) 900 ms, flip angle 9°, acquired resolution 0.68 x 0.68 x 1.60 mm, acquired matrix size 152 x 320 x 144, reconstructed resolution 0.34 x 0.34 x 1.60 mm (after two-fold interpolation in-plane by zero-filling in k-space), reconstructed matrix size 540 x 640 x 144, GRAPPA factor 2. Acquisition time: 5:25 min.

CPMG: Coronal, TR 4500 ms, TE 12 ms, number of echoes 10, echo spacing 12 ms, acquired resolution 0.68 x 0.68 x 1.7 mm inclusive of 15% slice gap, acquired matrix size 152 x 320, 34 slices, interleaved slice order, reconstructed resolution 0.34 x 0.34 x 1.7 mm (after two-fold interpolation in-plane by zero-filling in k-space, and inclusive of 15% slice gap), reconstructed matrix size 540 x 640, 34 slices, GRAPPA factor 2. Acquisition time: 11:07 min.

Study 2

This protocol has been previously described by Wearn et al. (2020a).

The imaging protocol included a 3D T1-weighted whole-brain MPRAGE and 2D multi-contrast turbo spin-echo (TSE).

MPRAGE: Sagittal, whole-brain, TR 2200 ms, TE 2.28 ms, TI 900 ms, flip angle 9°, FOV 220 x 220 x 179 mm, acquired resolution 0.86 x 0.86 x 0.86 mm, acquired matrix size 256 x 256 x 208. Acquisition time: 5:07 min.

Multi-contrast TSE: Coronal, TR 7500ms, number of echoes: 3, TE 9.1, 72 & 136 ms, acquired resolution 0.69 x 0.69 x 1.5 mm, reconstructed resolution 0.34 x 0.34 x 1.5 mm (after 2-fold interpolation in-plane by zero-filling in k-space, and inclusive of 15% slice gap), GRAPPA factor 2, FOV 220 x 220 x 34, acquired matrix size 270 x 320 x 58. Acquisition time: 5:09 min.

CPMG and TSE scans were not ‘whole-brain’, their coverage only extending approx. 1cm beyond anterior and posterior ends of the hippocampus. These scans were tilted such that the hippocampal body lay perpendicular to the slice acquisition plane.

The two distinct methods of measuring quantitative T2 (CPMG vs TSE) will give inherently different values for T2 midpoint and heterogeneity between studies (See supplementary information).

Relationships to variables such as age and cognitive score should be similar, given they are sensitive to the same tissue properties.

Imaging analyses

All analyses were performed at CRICBristol in a Linux cluster environment. All analyses were carried out in single-subject native space.

CPMG and TSE scans were brain-extracted using FSL's *bet2* on the first echo in the series (Smith, 2002). All extracted images were visually inspected for quality and rerun with different fractional intensity thresholds or gradient parameters where necessary. Fractional intensity threshold was typically set between 0.2-0.3. MPRAGE images were brain-extracted using *vbm8bet* (in-house script) and bias-field-corrected using FSL FAST (Zhang et al., 2001). T2 maps were created in MATLAB from multi-echo sequences by fitting logarithmic-space mono-exponential decay functions to each voxel series (overall summary of T2 calculation is shown in Knight et al. (2019)). The first echo of CPMG was always excluded. A sum-of-echoes image was created in order to have one structural image representing the entire multi-echo sequence. This image was used for segmentation.

Hippocampus was automatically masked using the Automatic Segmentation of Hippocampal Subfields (ASHS) software package (Yushkevich et al., 2015) (version: rev103, dated 12/06/2014; UPENN memory centre atlas dated 16/04/2014). This atlas contains a demographically similar template set to the present studies (healthy older adults and people with MCI). ASHS has demonstrated high accuracy whilst minimising subjective rater bias, without the need for group blinding (Example output shown in Figure 1). CA1, CA2 and CA3 were pooled to create a total "CA" mask, given the small size of CA2 and CA3. The outermost 1 voxel layer of each subfield was eroded before T2 histograms were calculated, in order to minimise the effects of partial voluming and extraneous brain areas confounding T2 histogram analyses. Pilot data (Knight et al., 2019) show

that using this method, extreme T2 values are not located around the outside border of the subfields, which would indicate strong partial voluming effects. Eroded subfield masks were overlaid onto T2 maps, giving a value of T2 for each voxel within each subfield.

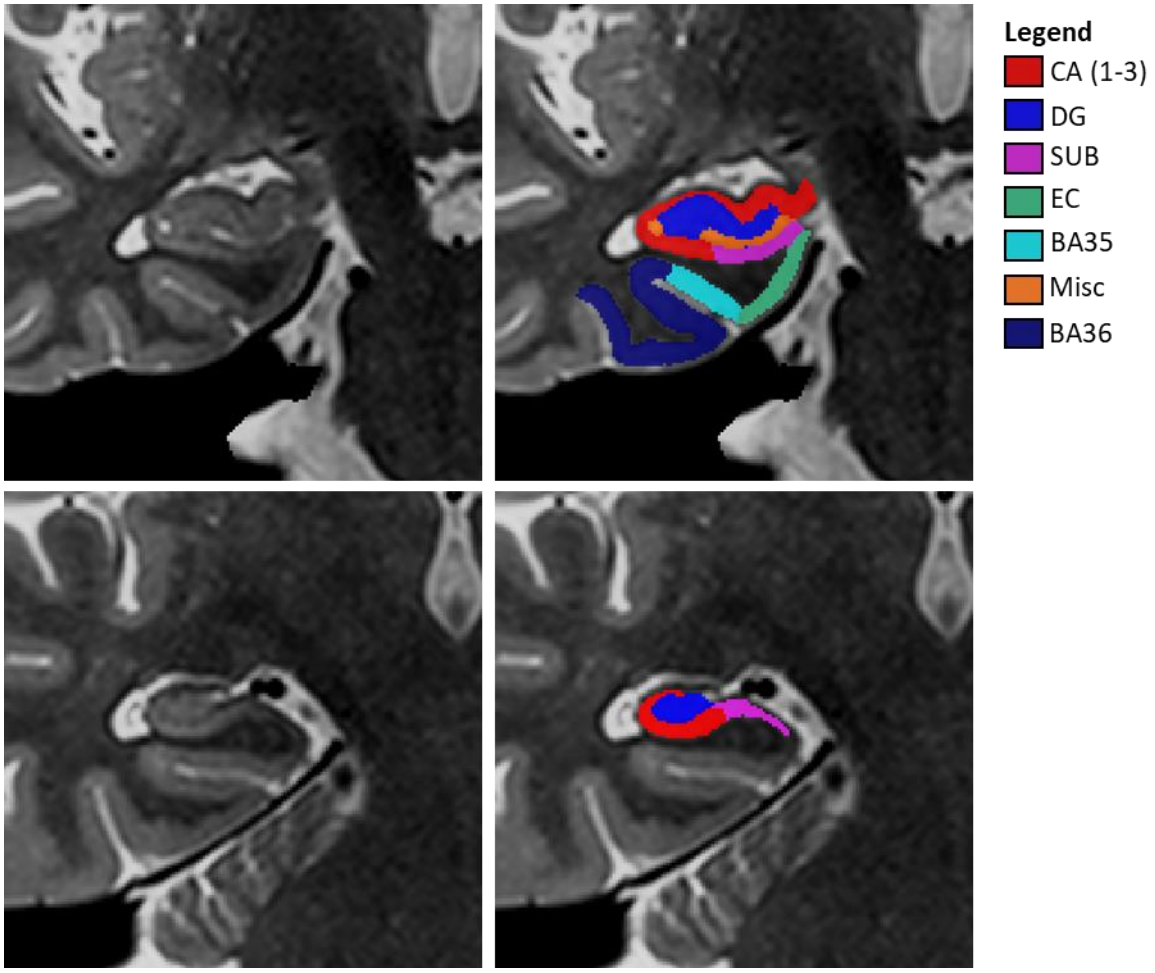

Figure 1 | ASHS MTL mask example with subfields labelled.
Each row shows unmasked (left) and masked (right) coronal slices of a healthy control MTL on summed-over-echoes T2-weighted scan. Top row shows hippocampal head, bottom row shows hippocampal body. DG = Dentate Gyrus, CA = Cornu Ammonis 1-3, SUB=Subiculum, EC = Entorhinal Cortex, BA = Brodmann Area. Misc and BA36 were excluded from all analyses. [COLOUR FIGURE]

Modelling T2 heterogeneity

T2 distribution histograms were modelled as loglogistic distributions within each subfield, as this was found to be the best fitting overall model in the whole hippocampus (Wearn et al., 2020a). Log-logistic distribution is defined as:

$$f\ (x|\mu,\sigma) = \frac{1}{\sigma}\frac{1}{x}\frac{\exp\ (z)}{[1+\exp\ (z)]^2}\ , \qquad where\ z = \frac{\log(x) - \mu}{\sigma}$$

where μ and σ denote the log-median value (midpoint) and distribution shape (heterogeneity), respectively. These descriptors of log-logistic distribution are analogous to median and median absolute deviation on a normal distribution. The models are described in more detail by Wearn et al. (2020a).

Statistical analysis

Volumes, T2 parameters and g eneral cognitive scores (MoCA / ACE-III) were pooled between studies after being converted into Z-scores for each study separately, with the left DG of healthy controls of each study as a reference point. In doing this, group, subfield, and hemisphere differences were maintained within each population.

We used a mixed model analysis, using the '*fitlme'* MATLAB function, to assess differences in T2 parameters between groups (HC, MCI), subfields (DG, CA, SUB, EC, BA35), hemispheres (left, right) and the effect of age. Full factorial models were created (assessing intercept and all possible interactions of aforementioned variables) with random effects of subject, subject*subfield interaction, and subject*hemisphere interaction. Each random effect significantly improved the model fit (as measured by AIC), without overparameterizing the model (confirmed by checking the hessian matrix and AIC). Age of the entire cohort was converted to a Z-score before being entered into the model. Models were created using Restricted Maximum Likelihood Estimation. Degrees of freedom were calculated using Satterthwaite approximation. A separate model was created for each of T2 midpoint (T2μ) and T2 heterogeneity (T2σ). Because of the inclusion of age in the models, all results are appropriately corrected for the effect of age.

Between-group *post hoc* comparisons were assessed using independent-samples t-tests. Cohen's *d* values for each comparison are reported as a measure of effect size. *Post hoc* effects of age in each subfield are assessed using linear regression.

Ability of volume and T2 in MTL subfields to predict cognitive decline (as measured by MoCA / ACE-III) was assessed using stepwise linear regression with forward selection. Forward stepwise regression works by first adding the variable (subfield) which best improves model fit (if any), then continuing to add variables which further significantly improve model fit. The criteria for adding more variables was a significant F statistic at the $p<0.05$ level. We entered follow-up cognition as the dependent variable and baseline cognition and age as fixed covariates:

$$\begin{aligned}\text{Follow-up Cognition} \\ = \beta_{intercept} + \beta_{age}(\text{Age}) + \beta_{BLCog}(\text{BaselineCognition}) + [\beta_{CA}(\text{CA}) + \beta_{DG}(\text{DG}) \\ + \beta_{SUB}(\text{SUB}) + \beta_{EC}(EC) + \beta_{BA35}(BA35) + \beta_{Hipp}(Hipp)] + \text{error}\end{aligned}$$

Parameters in square brackets represent those entered in a stepwise fashion. 'Hipp' represents total hippocampus. This analysis was performed for the MCI cohort only given the greater variability within this group. Z-scores for this analysis were calculated relative to each study's MCI population only. This method of predicting follow-up cognition whilst correcting for baseline cognition represents a more precise and less biased way of assessing cognitive change over time, compared to using change scores, as described by Vickers and Altman (2001).

We assessed the relationship between age, subfield T2 heterogeneity, subfield volume (ICV-corrected) and memory using path analysis. Memory was assessed using two measures of performance on the PAL task – total accuracy and mean reaction time across all trials. We took a semi-supervised approach when defining the model paths. The direction of the arrows was predetermined by theory; T2 heterogeneity is caused by microstructural changes that precede significant atrophy. In this regard, volumes were not allowed to predict T2. Memory scores were not allowed to predict structural variables, and, for obvious reasons, no variable was allowed to predict age. All direct effects in this direction were modelled and compared. Initially the model was run with both PAL scores in a single model, feeding into a single latent variable representing overall PAL score. However, we observed poor factor loadings of each score onto the latent variable (<.60), so

ran two separate models instead each with a single PAL outcome measure. Error terms were added to all variables except age and covariances and modification indices were calculated between all terms. All error term pairs within a structural measure (T2σ or volume) that had significant modification indices were allowed to covary, substantially improving overall model fit. In other words, subfields were allowed to covary within (but never between) each MRI modality. We observed poor model fit when data was entered as Z-scores normalised to left DG of HCs of each study as described above. We therefore normalised each structural measure to only the respective HC population, such that the mean ± standard deviation for each structural measure within any given subfield was 0 ± 1. In other words, differences between subfields were not considered by the model. This improved model fit to acceptable levels and should be considered in interpretation of the data.

Finally, we ran partial correlations to explore the relationship between absolute volume (uncorrected for intracranial volume) and T2 heterogeneity, correcting for age in healthy controls. The results of this analysis are presented in supplementary table 8.

All reported p-values are two-tailed. Where possible we have used comprehensive models, to minimise the need for multiple comparison corrections. False discovery rate (FDR) correction (Benjamini and Hochberg, 1995) was applied to individual paths and indirect paths in the path analysis, in line with guidance from Cribbie (2007). Data handling and storage was carried out using MathWorks MATLAB 2015a (with statistics and machine learning toolbox) and Microsoft Excel 2016. Mixed models were created and assessed in MathWorks MATLAB 2018a. Other statistical analyses were performed in IBM SPSS Statistics 24. Graphs were produced using GraphPad Prism v8.

RESULTS

Participant demographics

Demographic information for the cohort is displayed in Table 1. Our healthy control and MCI groups are closely matched for age and years of education. Details of cohorts in study 1 and study 2 are separately displayed in supplementary tables 1 and 2.

Table 1 | Participants demographics
*Cognitive score was calculated using MoCA (study 1) and ACE-III (study 2). Because of the different measures used between studies, each score was first normalised to each study's respective HC group, before being pooled here. Asterisks represent unpaired t-tests between groups (*p<0.05, ***p<.0001). YOE = Years of Education, HC = Healthy Control, MCI = Mild Cognitive Impairment.*

	HC	MCI	Total
N (male: female)	99 (47:52)	49 (27:22)	148 (74:74)
Age (years)	69.2 ± 8.55	72.2 ± 9.03	70.2 ± 8.79
YOE	15.8 ± 3.13	*14.2 ± 2.81	15.3 ± 3.11
Cognitive score (normalised to HC)	.000 ± 1.00	***-3.75 ± 2.42	-1.25 ± 2.39

Part 1: T2 changes in MCI

T2 heterogeneity

In addition to the overall effects of group ($F(1,144)=17.7$, $p<.0001$) and subfield ($F(4,576)=262$, $p<.0001$; explored further in supplementary information) and in line with hypothesis 1, the mixed model analysis revealed a significant group*subfield interaction ($F(4,576)=5.01$, $p=.001$). This effect did not vary according to hemisphere so a two-way ANOVA on pooled hemispheres was conducted with predicted values from the mixed model analysis, with pairwise group comparisons for each subfield (Figure 2). This test revealed a significantly greater T2 heterogeneity in the MCI group in all subfields but with varying effect sizes (DG: $t=3.26$, $p_{corr}=.008$, Cohen's $d=0.59$; CA: $t=3.79$, $p_{corr}=.001$, Cohen's $d=0.68$; SUB: $t=5.50$, $p_{corr}<.0001$, Cohen's $d=0.99$; EC: $t=6.63$, $p_{corr}<.0001$, Cohen's $d=1.22$; BA35: $t=5.26$, $p_{corr}<.0001$, Cohen's $d=0.96$, all adjusted using Bonferroni correction for multiple comparisons).

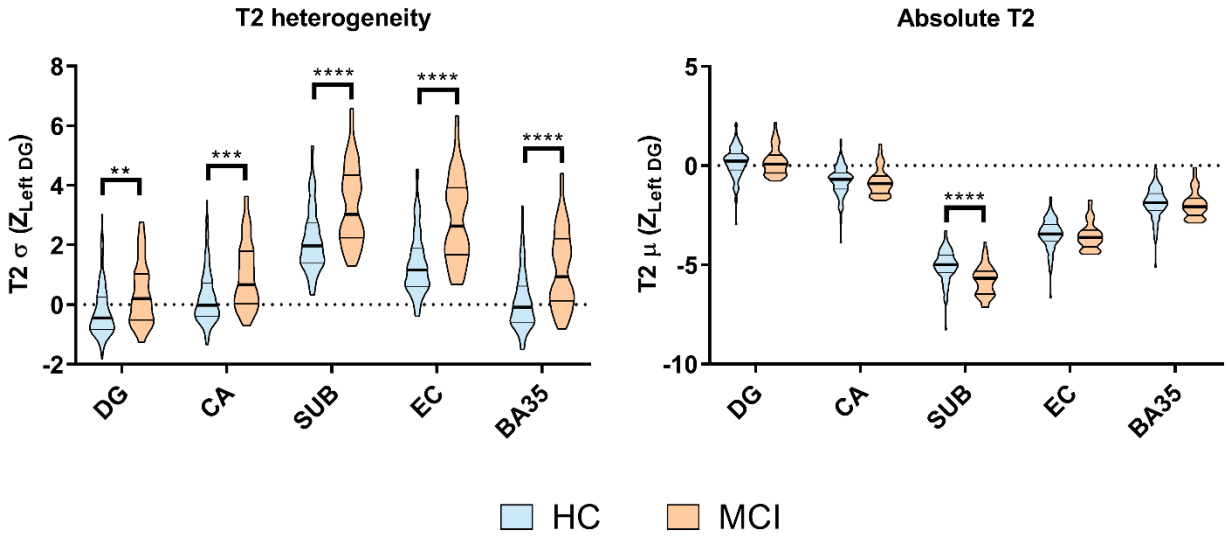

Figure 2 | Subfield comparisons for T2 heterogeneity and absolute T2.
*Violin plots showing group & subfield differences (pooled across hemispheres & corrected for age) for absolute T2 and heterogeneity marginal means. HC = healthy older control; MCI = mild cognitive impairment. Stars represent p-values from post hoc two-way ANOVA tests to compare subfield groupwise differences. **p<.01; ***p<.001; ****p<.0001. [COLOUR FIGURE]*

We also found a significant main effect of age ($F(1,144)=37.6$, $p<.0001$) and an interaction between subfield and age ($F(4,576)=10.4$, $p<.0001$). Older people had significantly increased T2 heterogeneity in all subfields. The association increased in strength in a stepwise fashion through DG ($R^2=.106$, $F(1,146)=17.3$, $p_{corr}=.0003$), CA ($R^2=.147$, $F(1,146)=27.5$, $p_{corr}<.0001$), SUB ($R^2=.159$, $F(1,146)=27.5$, $p_{corr}<.0001$), EC ($R^2=.185$, $F(1,146)=33.2$, $p_{corr}<.0001$) and BA35 ($R^2=.227$, $F(1,146)=42.8$, $p_{corr}<.0001$).

Finally, we observed no interactions between group and age ($F(1,144)=1.74$, $p=.190$) or group and hemisphere ($F(1,144)=0.31$, $p=.579$) nor were there any three- or four-way interactions for T2 heterogeneity.

Absolute T2

For comparison, we also explored subfield-specific changes in absolute T2. In contrast to T2 heterogeneity we observed no overall difference between HC and MCI groups on absolute T2

(F(1,144)=1.19, p=.278), but did see a substantial overall difference in absolute T2 between subfields (F(4,576)=974, p<.0001; explored further in supplementary information). We found a group*subfield interaction (F(4,576)=3.78, p=.005) driven by a subiculum-specific low T2μ in the MCI group (t=5.50, p_{corr}<.0001, Cohen's *d*=0.95). In no other subfield was there any difference between groups.

Although we observed no significant main effect of age on absolute T2 (F(1,144)=0.01, p=.907) , the model did reveal a significant interaction between subfield and age (F(4,576)=2.92, p=.021). However, in no subfield was there any statistically significant association between age and absolute T2 (all subfields: p > .250).

We observed no interactions between group and age (F(1,144)=.005, p=.946) or group and hemisphere (F(1,144)=1.65, p=.201) nor were there any three- or four-way interactions for absolute T2.

Predicting cognitive change over time

In order to identify whether MTL subfields could predict cognitive decline, we ran three stepwise linear regressions to predict cognitive score of people with MCI after one year. One model was run for each MRI modality: volume, T2 heterogeneity, absolute T2, whilst always accounting for baseline cognitive score and age. Age and baseline cognition alone were unable to significantly predict follow-up cognitive score (R^2=.092, F(2,19)=.859, p=.441; AIC=29.9). In other words, age alone was unable to predict the degree of cognitive change over the year.

Of all three modalities, only T2 heterogeneity could accuracy predict cognitive change over time. In this model, greater T2 heterogeneity within EC predicted poorer cognition after the year (R^2=.406, F(3,19)=3.65, p=.035, AIC=23.4; β_{EC}=-.665, p=.010). No additional subfields sufficiently increased the model fit to qualify for entry into the model.

EC volume was entered as sole predictor of greater cognitive decline in the volume model (β_{EC}=.595, p=.030). However the full model was not statistically significant (R^2=.329, F(3,19)=2.62, p=.087, AIC=25.8), indicating a weak association between subfield volume and cognitive change. Similarly, absolute T2 in DG was selected as a significant predictor of cognitive score after one year, with higher values predicting greater cognitive decline (β_{DG}=-.494, p=.044), however the overall absolute T2 model was not statistically significant (R^2=.301, F(3,19)=2.30, p=.117, AIC=26.6).

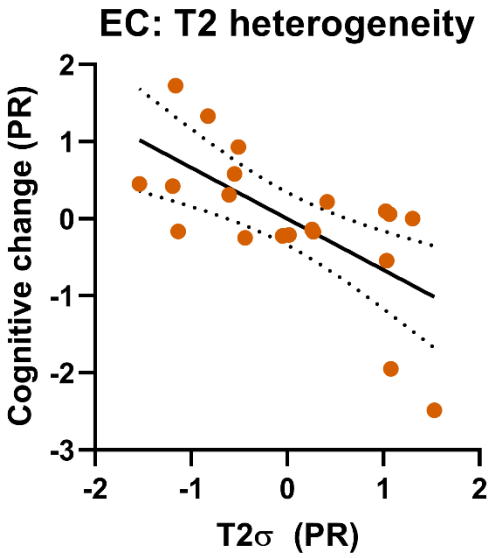

Figure 3 | Subfield structure predicting cognitive decline in people with MCI
Partial residual (PR) plot of significant model from stepwise linear regression. Graph shows linear regression lines (solid lines) ± 95% confidence intervals (dotted lines). The only significantly predictive model created was that involving T2 heterogeneity within entorhinal cortex (EC). [COLOUR FIGURE]

Part 2: Sequential relationships across brain structure and behaviour

We used path analysis to explore predicted relationships between age, T2 heterogeneity, subfield volume (relative to intracranial volume) and cognition. We took a semi-supervised approach when defining the model paths. The direction of the arrows was predetermined on the theoretical basis that T2 heterogeneity is likely caused by microstructural changes that precede significant atrophy. Memory scores were outcome variables, so were not allowed to predict structural variables, and no variable was allowed to predict age. All direct effects in this direction were modelled and compared.

Two separate models were created, one for each of the PAL output measures (See Methods). After applying covariance structures to error terms between subfields, model fit was good with various fit parameters falling within an acceptable range (χ^2=21.8, df = 20, p=.350; C_{min}/DF=1.09; GFI=.964, AGFI=.861; CFI=.998; RMSEA=.031 ± 90% CI [<.001-.096]; p_{close}=.615). Model fit was identical between the two models. Covariance matrices can be found in supplementary information (Supplementary Tables 4 & 5).

The final path analysis models (Figure 4) reveal the following relationships, for which statistics are shown in Table 2 and Supplementary Table 3. Age is a significant positive predictor of T2 heterogeneity in all MTL subfields. In turn, T2 heterogeneity within each subfield is a significant negative predictor of subfield volume. The only significant direct effects of age on subfield volume are seen with DG and CA (where greater age predicts smaller volumes), however, significant indirect effects are seen between age and subfield volume within all subfields (Table 2). T2 heterogeneity therefore at least partially mediates the relationship between age and volume.

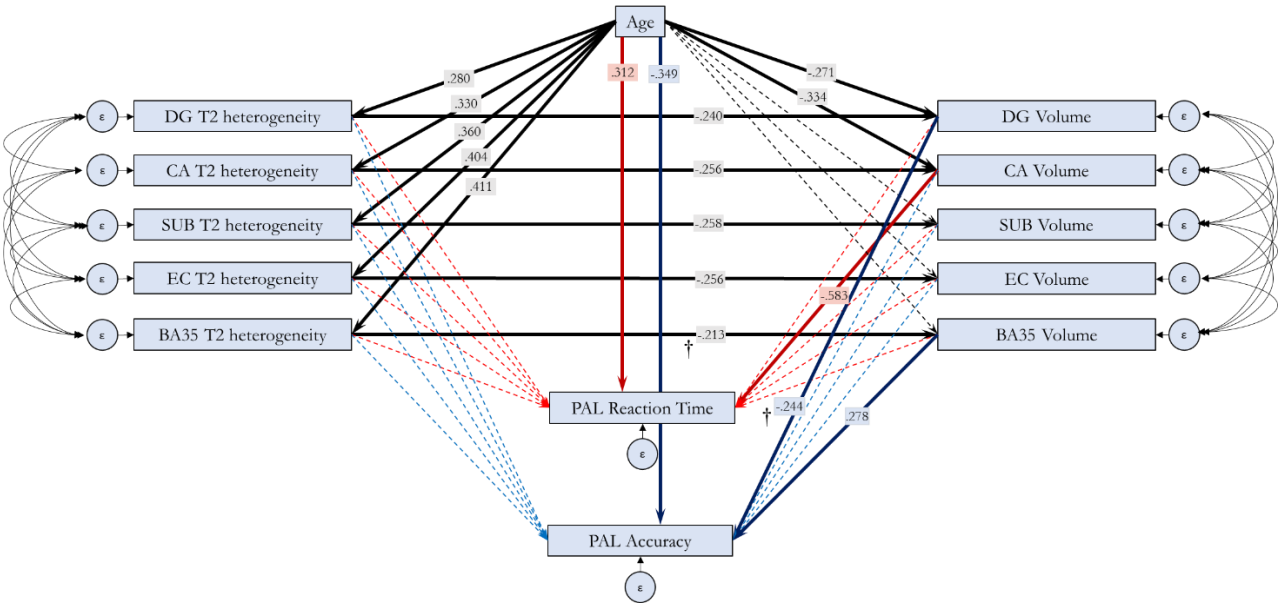

Figure 4 | Path analysis showing the relationship between Age, T2 heterogeneity, volume and memory in MTL subfields in healthy older controls.
Bold arrows represent statistically significant relationships (p<.05), with standardized B values indicated in overlaid boxes. Paths that do not survive FDR correction for multiple comparisons are marked with †. Two models were run, each assessing one outcome measure of the PAL task. Black lines represent paths shared between the models. Unique paths to each model are shown in red (PAL reaction time as dependent variable) and blue (PAL Total Accuracy as dependent variable). Curves lines represent error term

covariances defined in the model. All subfield volumes were normalised to ICV prior to entering into the model. Full statistics are shown in Supplementary Table 3. [COLOUR FIGURE]

No significant direct effects were observed between T2 heterogeneity and either PAL score (Table 2). However, a direct negative effect of CA volume on PAL mean reaction time was observed, as well as a direct positive effect of BA35 volume on PAL accuracy. In contrast, we also see that DG volume has a direct negative relationship with PAL accuracy. The indirect relationship between T2 heterogeneity for each of these respective subfields and each respective PAL score is statistically significant (Table 2), indicating a role of volume as a mediator between T2 heterogeneity and cognition. Interestingly, analysis of indirect paths also reveals a marginally significant positive relationship between DG T2 heterogeneity and PAL reaction time.

Finally, outside of MTL subfield structure and microstructure, we also observe direct effects of age on PAL scores, with greater age predicting poorer scores on both measures (Table 2).

This all serves to show a clear pathway of the effect of age on cognition, first through its increasing of T2 heterogeneity in subfields, which in turn reduces the volume of each subfield, which then, in a subfield-specific manner, reduces cognitive functionality. The full indirect pathway from age, through T2 heterogeneity, to volume and finally to PAL score is statistically significant for pathways through DG and BA35 for predicting PAL accuracy, and through CA for predicting PAL reaction time (Table 2). Outside any MTL subfield paths, there remains a direct effect of age on PAL score, whereby greater age predicts poorer performance (lower accuracy, longer reaction times), suggesting that age affects PAL performance through non-hippocampal as well as hippocampal pathways.

We ran the models again on the MCI group data to explore any differences in path structure to the HC models. We observed no major changes between the models other than the loss of the direct relationship between age and PAL scores. We also ran a model with the volumes allowed to predict T2 heterogeneity. this reverse path analysis shows that T2σ does not have the same mediatory effect

414 on volume that volume has on T2σ and has poorer model fit parameters. These models will not be
415 discussed further here as they are not the focus of this study, but the model path diagrams can be
416 seen in supplementary information (supplementary figures 1 & 2 and supplementary tables 4 & 5).
417

418 ***Table 2 | Summary statistics of key indirect paths***
419 *All paths remain statistically significant after FDR correction for multiple comparisons (corrected p-value threshold = .0203)*

Indirect path		Standardized estimate	P-value
Age → T2 heterogeneity → Volume			
	DG	-.072	*.002
	CA	-.079	*.002
	SUB	-.093	*.002
	EC	-.103	*.005
	BA35	-.087	*.009
T2 heterogeneity → Volume → PAL			
PAL Accuracy	DG	.063	*.020
	CA	-.044	.175
	SUB	-.032	.177
	EC	.008	.753
	BA35	-.059	*.013
PAL Reaction Time	DG	-.050	.130
	CA	.140	*.002
	SUB	-.040	.173
	EC	-.009	.592
	BA35	.021	.220
Age → T2 heterogeneity → Volume → PAL			
PAL Accuracy	DG	.016	*.013
	CA	-.016	.135
	SUB	-.012	.131
	EC	.003	.749
	BA35	-.024	*.011
	Total path (all subfields)	-.006	.183
PAL Reaction Time	DG	-.013	.096
	CA	.049	*.002
	SUB	.056	.137
	EC	-.004	.585

BA35	.009	.183
Total path (all subfields)	.019	.174

DISCUSSION

Here we have shown that T2 heterogeneity in the MTL increases with cognitive impairment in a subfield-dependent manner in line with progression patterns of Alzheimer's pathology. We also show that T2 heterogeneity in the entorhinal cortex predicts cognitive decline in people with MCI where volume or T2 changes in any other subfield cannot. Furthermore, we describe a mechanism by which cognitive ability deteriorates with age through direct effects on T2 heterogeneity, which lead to the changes in volume which in turn lead to cognitive decline. These findings are discussed in detail below.

Subfield-specific differences in T2 heterogeneity in MCI and with age

T2 heterogeneity was greater in MCI across all subfields of the MTL, however the magnitude of this effect differed between subfields. The smallest increase was seen in DG, and the largest in EC and BA35. This is in line with literature on the timing of deposition of NFTs throughout the course of AD (Braak and Braak, 1991, Braak and Braak, 1995), whereby the transentorhinal region (comprising BA35 and lateral EC) is affected first, and DG pathology is only detectible later on. DG is also one of the last MTL regions to exhibit volume loss (Daugherty et al., 2015). This accordance between the pattern of T2 heterogeneity differences and AD neuropathological progression supports T2 heterogeneity as a means to detect pathologically relevant change before the onset of dementia.

This is in contrast to absolute T2, which is no different between healthy controls and people with MCI in any subfield other than subiculum (where it is decreased in MCI). The effect may be explained by an increase in iron and compact amyloid plaques, which have been seen to be increased

in SUB and CA1 regions compared to DG and CA3 in a mouse model of AD (Reilly et al., 2003). Even though we cannot directly show the prevalence of AD pathology in our MCI sample, these results are in line with this group being at higher risk for displaying AD pathology.

Our findings support our proposal (Knight et al., 2019, Wearn et al., 2020a) that T2 heterogeneity defines hippocampal integrity at the subfield level better than absolute T2. In further support of this model, two studies of absolute T2 in hippocampal subfields *in vitro* (Huesgen et al., 1993, Antharam et al., 2012) note a lack of any consistent pattern between healthy control and Alzheimer's slices in any subfield.

Clinical Utility of T2 heterogeneity in MTL subfields

We show that T2 heterogeneity in EC could predict cognitive decline in our MCI cohort, where volume or absolute T2 in whole hippocampus or other MTL subfields could not. Measuring T2 heterogeneity in EC may therefore provide the best chance to identify those who have MCI due to incipient Alzheimer's disease as opposed to other causes, though further testing is required to confirm this.

T2 heterogeneity in the EC possessed the greatest power for predicting cognitive decline of the subfields tested. EC structure has repeatedly been shown to better predict conversion to Alzheimer's disease than hippocampal structure (most often measured by volume or thickness, reviewed by Zhou et al. (2016)). However, it is interesting to note that EC is selected as a better predictor over BA35, despite BA35 supposedly being an earlier NFT deposition site in AD. One reason could be that the transentorhinal cortex spans both regions, and that subsequent microstructural changes are more widespread across EC compared to BA35 in MCI, supported by histological data (Braak and Tredici, 2015). It is important to note that other subfields were unable to improve model fit as a likely result of shared variance being explained by each subfield. In other words, we cannot conclude that the predictive power provided by EC was 'significantly greater' than that provided by other subfields.

Rather, other subfields did not provide a significant amount of additional information on top of that provided by heterogeneity in EC.

The use of T2 heterogeneity is highly translatable to clinical settings. MRI is standard practice in improving the accuracy of a diagnosis of AD. The MRI scan necessary to calculate quantitative T2 can be completed within a few additional minutes of a standard clinical MRI (a multi-echo T2 sequence of sufficient resolution is all that is required). The same high-resolution scan can be used to automatically segment subfields of the MTL e.g. using ASHS (Yushkevich et al., 2015). Although the use of MRI in people with dementia can sometimes be tricky due exacerbated feelings of claustrophobia and confusion, this MRI protocol would be most useful in prodromal stages of the disease, before significant MTL volume loss and symptom severity, minimizing these complications. A cheaper theoretical screening test with high sensitivity for detecting pathology such as has been shown with tests of accelerated forgetting on word list tasks (Wearn et al., 2020b, Weston et al., 2018) or a blood test (Toombs and Zetterberg, 2020) may identify individuals who should qualify for an MRI scan sensitive to very early pathological hallmarks.

Our method of measuring heterogeneity of an MRI signal is arguably a form of texture analysis, a technique for detecting microstructural changes on MRI whose clinical applications are increasingly a topic of interest (for a review see Cai et al. (2020)). Zhao et al. (2020) highlight the use of 'radiomic biomarkers' (a method of texture analysis) as a robust method of predicting longitudinal change, and genetic risk for Alzheimer's disease. Similarly, Sørensen et al. (2017) find that various measures of hippocampal structure including volume and texture can accurately discriminate healthy older people from those with MCI and Alzheimer's disease. To our knowledge, no study has attempted texture analysis using quantitative T2, or in MTL subfields, making this study the first of its kind. This appears to be an emerging field with increasing potential for scientific interest and clinical application.

Although we have focussed on T2 changes due to Alzheimer's pathology, T2 heterogeneity is a novel measure which could be applied to any neurological disease characterised by microstructural changes, including other dementias, acute stroke (as demonstrated by Norton et al. (2017)), epilepsy or schizophrenia. Future research could utilise T2 heterogeneity to easily probe microstructural abnormalities. This study builds on our past analyses of T2 heterogeneity by highlighting how it can reveal structural changes on an even finer scale, indicating its usefulness in disorders where brain damage is highly localised.

Subfield-specific contributions to memory

We aimed to better understand the relationship between T2 heterogeneity, volume, age and cognitive ability in cognitively healthy older people. To summarise, our path analysis revealed the following pattern of relationships. T2 heterogeneity mediates the negative relationship between age and volume. Volume in turn mediates the negative relationship between T2 heterogeneity and memory, in a subfield-specific manner. In line with hypotheses 3 and 4, increased T2 heterogeneity therefore seems to represent a state of structural damage which may give rise to the volumetric changes which have the most profound impact on memory and cognition. Interestingly, we do still find a direct association between age and CA and DG volume, indicating only partial mediation by T2 heterogeneity in these regions. This suggests that there are other age-mediated volumetric changes which are not accurately reflected by changes in T2 heterogeneity.

Our results reveal potential subfield-specific associations with memory scores unlike previous similar models that have looked at whole hippocampus only (Rodrigue et al., 2012). Longer reaction times were associated with a subtle damage to the CA region, whereas poor accuracy was associated with low BA35 integrity but a more intact DG. This latter result was surprising but may suggest that an impaired BA35 *relative to* DG is indicative of a network prone to specific deficits in total accuracy. This pattern reflects changes seen in early AD, as the transentorhinal cortex is the first

region to display NFT deposition, and DG is the last of all regions tested (Braak and Braak, 1995, Braak and Braak, 1991, Xie et al., 2018). In other words, patterns of very subtle neurodegeneration akin to those found in the earliest stages of AD may give rise to specific deficits in total accuracy scores on the CANTAB PAL. It should be noted that this explanation is speculative and is subject to confirmation in other cohorts. Indeed, the relationship between DG volume and PAL accuracy does not appear to survive FDR correction for multiple comparisons and may therefore simply be an artifact of the modelling procedure. Nonetheless it indicates a point of interest for future studies to examine hippocampal subfield specificity to behavioural measures, and to further explore to what degree individual subfield structural changes may indicate early signs of pathology.

Limitations

The main limitation is the lack of biomarker status availability for people with MCI in this study (assessed either from Positron Emission Tomography or CSF analysis). Therefore, we cannot be certain as to the exact proportion of those who have incipient dementia. We suspect that T2 heterogeneity in entorhinal cortex would identify those who do have incipient Alzheimer's disease, as EC is one of the earliest sites of pathological change. This is supported by the ability of EC T2 heterogeneity in predicting cognitive decline. However, co-pathologies are almost certainly present in the overall cohort which may include other dementias or undiagnosed microbleeds.

Additionally, both studies utilised different scanning sequences for assessing quantitative T2, giving inherently different absolute T2 values. For this reason, we were careful to normalise results from each cohort before combining them. Although a potential confounding factor, we have noted in our previous study that effects of T2 heterogeneity do not appear to be specific to a sequence (Wearn et al., 2020a). Rather, we provide evidence that the different sequences are sensitive to the same physiological changes. This supports the translatability of this measure to clinical settings, where available sequences for measuring quantitative T2 may vary between sites.

We also note a potential selection bias in our 1-year follow-up analysis, as we experienced a relatively high dropout rate of our MCI cohort. Although reasons for not returning were never formally quantified, this was occasionally due to the participant feeling unable to attend the second session due to significant cognitive and/or functional decline. Those who declined the most may therefore be simply missed out of the current analyses. We expect the result of this bias to minimise the ability of our MRI variables to predict decline, so it is testament to the clinical potential of T2 heterogeneity that it can still predict cognitive decline in this cohort.

Finally, we acknowledge that our masking procedure of hippocampal subfields (automated using ASHS) has only been verified for measuring shape and volume, not T2 heterogeneity. Measures could feasibly be biased due to the inclusion of extraneous brain regions or subject to error from surrounding regions through partial voluming. We have minimised the risk of these factors by eroding our subfield masks by their outermost voxels. Furthermore, we find no evidence of a relationship between absolute subfield volume and T2 heterogeneity, indicating that the two measures are distinct, and not confounded by one another (supplementary information section 4).

Conclusions

The analyses presented in this paper comprise the first detailed exploration of quantitative T2 across subfields of the medial temporal lobe in older people with and without cognitive impairment. We support previous evidence that absolute T2 is not a sensitive marker of early pathology, rather, heterogeneity of T2 is much more sensitive to early pathological change. We demonstrate that T2 heterogeneity differs between subfields in a manner which reflects the order of NFT deposition in prodromal AD. We provide evidence that although T2 heterogeneity increases with age in all subfields, the degree to which this occurs is subfield-dependent and is strongest in MTL cortical regions (EC and BA35). In contrast, we do not see systematic evidence of a relationship between age and absolute T2. Using path analysis, we describe a pathway through which cognition is significantly

affected by age through direct effects on T2 heterogeneity in cognitively healthy older people, which in turn has direct effects on volume which lead to changes in cognition, supporting the idea that T2 changes precede and lead to volumetric changes. Finally, we show that the single best predictor of cognitive decline in people with MCI is greater T2 heterogeneity within entorhinal cortex, outperforming both T2 midpoint and, critically, volume of any subfield.

Declarations

Ethics approval and consent to participate

All patients provided informed written consent prior to testing. Ethical approval was given by Frenchay NHS Research Ethics Committee.

Consent for publication

Not applicable

Data and code availability

The datasets used during the current study are available from the corresponding author on reasonable request and setup of a formal data sharing agreement. Code used to compute T2 maps is available to download from https://data.bris.ac.uk/data/dataset/1bjytiabmtwqx2kodgbzkwso0k (Wearn et al., 2017)

Competing interests

We declare that none of the authors have competing financial or non-financial interests.

Funding

This research was funded in part by the Wellcome Trust [Grant number 109067/Z/15/AI]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. This study also was funded by Alzheimer's Research UK and BRACE.

Authors' contributions

Alfie Wearn: Conceptualisation, data curation, formal analysis, funding acquisition, investigation, Methodology, project administration, resources, software, visualisation, writing – original draft, writing – review & editing; Volkan Nurdal: Data curation, investigation, resources; Esther Saunders-Jennings: Data curation, investigation, resources; Michael J. Knight: Data curation,

resources, software, writing – review & editing; Christopher R. Madan: Formal analysis, methodology, writing – review & editing; Sean-James Fallon: Formal analysis, methodology, writing – review & editing; Hanna K. Isotalus: Investigation, resources; Risto A. Kauppinen: Funding acquisition, methodology, supervision, validation, writing – review & editing; Elizabeth J. Coulthard: conceptualisation, funding acquisition, methodology, project administration, resources, supervision, validation, writing – review & editing.

Acknowledgments

The authors wish to thank Join Dementia Research and the Avon & Wiltshire Mental Health Partnership for their assistance with participant recruitment. We also wish to thank those who have helped collect data for the projects (Serena Dillon, Demitra Tsivos, Emma Hadley, Ellen Gaaikema, Lucy Adams, Candida Stainer, Ben Kershaw & Bryony McCann), Aileen Wilson for her help conducting MRI scans, and all the volunteers who gave up their time to take part in our studies.

REFERENCES

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., WOLK, D. A. & 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. *Proceedings of the National Academy of Sciences,* 115**,** 201801093.

ALBERT, M. S., DEKOSKY, S. T., DICKSON, D., DUBOIS, B., FELDMAN, H. H., FOX, N. C., GAMST, A., HOLTZMAN, D. M., JAGUST, W. J., PETERSEN, R. C., SNYDER, P. J., CARRILLO, M. C., THIES, B. & PHELPS, C. H. 2011. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia,* 7**,** 270-279.

ALZHEIMER'S ASSOCIATION 2015. 2015 Alzheimer's disease facts and figures. *Alzheimer's & Dementia,* 11**,** 332-384.

ANTHARAM, V., COLLINGWOOD, J. F., BULLIVANT, J.-P., DAVIDSON, M. R., CHANDRA, S., MIKHAYLOVA, A., FINNEGAN, M. E., BATICH, C., FORDER, J. R. & DOBSON, J. 2012. High field magnetic resonance microscopy of the human hippocampus in Alzheimer's disease: Quantitative imaging and correlation with iron. *NeuroImage,* 59**,** 1249-1260.

APOSTOLOVA, L. G., DUTTON, R. A., DINOV, I. D., HAYASHI, K. M., TOGA, A. W., CUMMINGS, J. L. & THOMPSON, P. M. 2006. Conversion of Mild Cognitive Impairment to Alzheimer Disease Predicted by Hippocampal Atrophy Maps. *Archives of Neurology,* 63**,** 693-699.

APOSTOLOVA, L. G., MOSCONI, L., THOMPSON, P. M., GREEN, A. E., HWANG, K. S., RAMIREZ, A., MISTUR, R., TSUI, W. H., DE LEON, M. J., APOSTOLOVA, L. G., MOSCONI, L., THOMPSON, P. M., GREEN, A. E., HWANG, K. S., RAMIREZ, A., MISTUR, R., TSUI, W. H. & DE LEON, M. J. 2010. Subregional hippocampal atrophy predicts Alzheimer's dementia in the cognitively normal. *Neurobiology of Aging,* 31**,** 1077-88.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society.* , 57**,** 289-300.

BRAAK, H. & BRAAK, E. 1991. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica,* 82**,** 239-259.

BRAAK, H. & BRAAK, E. 1995. Staging of alzheimer's disease-related neurofibrillary changes. *Neurobiology of Aging,* 16**,** 271278.

BRAAK, H. & TREDICI, K. D. 2015. The preclinical phase of the pathological process underlying sporadic Alzheimer's disease. *Brain,* 138**,** 2814-2833.

BRIELLMANN, R. S., KALNINS, R. M., BERKOVIC, S. F. & JACKSON, G. D. 2002. Hippocampal pathology in refractory temporal lobe epilepsy. *Neurology,* 58**,** 265-271.

CAI, J.-H., HE, Y., ZHONG, X.-L., LEI, H., WANG, F., LUO, G.-H., ZHAO, H. & LIU, J.-C. 2020. Magnetic Resonance Texture Analysis in Alzheimer's disease. *Academic Radiology,* 27**,** 1774-1783.

CHÉTELAT, G., FOUQUET, M., KALPOUZOS, G., DENGHIEN, I., LA SAYETTE, D. V., VIADER, F., MÉZENGE, F., LANDEAU, B., BARON, J. C., EUSTACHE, F., DESGRANGES, B., CHÉTELAT, G., FOUQUET, M., KALPOUZOS, G., DENGHIEN, I., LA SAYETTE, D. 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**,** 1721-31.

CRIBBIE, R. A. 2007. Multiplicity Control in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal,* 14**,** 98-112.
CUMMINGS, J. L., MORSTORF, T. & ZHONG, K. 2014. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimer's Research & Therapy,* 6**,** 37.
DAUGHERTY, A. M., BENDER, A. R., RAZ, N. & OFEN, N. 2015. Age differences in hippocampal subfield volumes from childhood to late adulthood. *Hippocampus,* 26**,** 220-8.
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. *Human Brain Mapping,* 36**,** 463-474.
DETOLEDO-MORRELL, L., STOUB, T. R., BULGAKOVA, M., WILSON, R. S., BENNETT, D. A., LEURGANS, S., WUU, J. & TURNER, D. A. 2004. MRI-derived entorhinal volume is a good predictor of conversion from MCI to AD. *Neurobiology of Aging,* 25**,** 1197-1203.
DUVERNOY, H. M., CATTIN, F., RISOLD, P. Y., VANNSON, J. L. & GAUDRON, M. 2013. *The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI*, Springer Berlin Heidelberg.
ÉGERHÁZI, A., BERECZ, R., BARTÓK, E. & DEGRELL, I. 2007. Automated Neuropsychological Test Battery (CANTAB) in mild cognitive impairment and in Alzheimer's disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry,* 31**,** 746-751.
FRISONI, G. B., GANZOLA, R., CANU, E., RÜB, U., PIZZINI, F. B., ALESSANDRINI, F., ZOCCATELLI, G., BELTRAMELLO, A., CALTAGIRONE, C., THOMPSON, P. M., 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,* 131**,** 3266-76.
FRISONI, G. B., SABATTOLI, F., LEE, A. D., DUTTON, R. A., TOGA, A. W., THOMPSON, P. M., FRISONI, G. B., SABATTOLI, F., LEE, A. D., DUTTON, R. A., TOGA, A. W. & THOMPSON, P. M. 2006. In vivo neuropathology of the hippocampal formation in AD: A radial mapping MR-based study. *NeuroImage,* 32**,** 104-10.
FUKUTANI, Y., CAIRNS, N. J., SHIOZAWA, M., SASAKI, K., SUDO, S., ISAKI, K. & LANTOS, P. L. 2000. Neuronal loss and neurofibrillary degeneration in the hippocampal cortex in late‐onset sporadic Alzheimer's disease. *Psychiatry and Clinical Neurosciences,* 54**,** 523-529.
FUKUTANI, Y., KOBAYASHI, K., NAKAMURA, I., WATANABE, K., ISAKI, K. & CAIRNS, N. J. 1995. Neurons, intracellular and extracellular neurofibrillary tangles in subdivisions of the hippocampal cortex in normal ageing and Alzheimer's disease. *Neuroscience Letters,* 200**,** 57-60.
HOUSE, M. J., PIERRE, T. G. S. & MCLEAN, C. 2008. 1.4T study of proton magnetic relaxation rates, iron concentrations, and plaque burden in Alzheimer's disease and control postmortem brain tissue. *Magnetic Resonance in Medicine,* 60**,** 41-52.
HUESGEN, C. T., BURGER, P. C., CRAIN, B. J. & JOHNSON, G. A. 1993. In vitro MR microscopy of the hippocampus in Alzheimer's disease. *Neurology,* 43**,** 145-145.
INGELSSON, M., FUKUMOTO, H., NEWELL, K. L., GROWDON, J. H., HEDLEY–WHYTE, E. T., FROSCH, M. P., ALBERT, M. S., HYMAN, B. T. & IRIZARRY, M. C. 2004. Early A-beta accumulation and progressive synaptic loss, gliosis, and tangle formation in AD brain. *Neurology,* 62**,** 925-931.
JACK, C. R., KNOPMAN, D. S., JAGUST, W. J., SHAW, L. M., AISEN, P. S., WEINER, M. W., PETERSEN, R. C. & TROJANOWSKI, J. Q. 2010. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology,* 9**,** 119-128.

JACK, C. R., SLOMKOWSKI, M., GRACON, S., HOOVER, T. M., FELMLEE, J. P., STEWART, K., XU, Y., SHIUNG, M., O'BRIEN, P. C., CHA, R., KNOPMAN, D. & PETERSEN, R. C. 2003. MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD. *Neurology,* 60**,** 253-260-260.

KERCHNER, G. A., DEUTSCH, G. K., ZEINEH, M., DOUGHERTY, R. F., SARANATHAN, M. & RUTT, B. K. 2012. Hippocampal CA1 apical neuropil atrophy and memory performance in Alzheimer's disease. *NeuroImage,* 63**,** 194-202.

KNIGHT, M. J., WEARN, A., COULTHARD, E. & KAUPPINEN, R. A. 2019. T2 Relaxometry and Diffusion Tensor Indices of the Hippocampus and Entorhinal Cortex Improve Sensitivity and Specificity of MRI to Detect Amnestic Mild Cognitive Impairment and Alzheimer's Disease Dementia. *Journal of Magnetic Resonance Imaging,* 49**,** 445-455.

LA JOIE, R., PERROTIN, A., DE SAYETTE, V., EGRET, S., DOEUVRE, L., BELLIARD, S., EUSTACHE, F., DESGRANGES, B., CHÉTELAT, G., JOIE, R., PERROTIN, A., DE SAYETTE, V., EGRET, S., DOEUVRE, L., BELLIARD, S., EUSTACHE, F., DESGRANGES, B. & CHÉTELAT, G. 2013. Hippocampal subfield volumetry in mild cognitive impairment, Alzheimer's disease and semantic dementia. *NeuroImage: Clinical,* 3**,** 155-62.

LAAKSO, M. P., PARTANEN, K., SOININEN, H., LEHTOVIRTA, M., HALLIKAINEN, M., HÄNNINEN, T., HELKALA, E. L., VAINIO, P. & RIEKKINEN, P. J. 1996. MR T2 relaxometry in Alzheimer's disease and age-associated memory impairment. *Neurobiology of Aging,* 17**,** 535-540.

LEE, D., THALER, J. P., BERKSETH, K. E., MELHORN, S. J., SCHWARTZ, M. W. & SCHUR, E. A. 2013. Longer T2 relaxation time is a marker of hypothalamic gliosis in mice with diet-induced obesity. *American Journal of Physiology-Endocrinology and Metabolism,* 304**,** E1245-E1250.

MEADOWCROFT, M. D., PETERS, D. G., DEWAL, R. P., CONNOR, J. R. & YANG, Q. X. 2015. The effect of iron in MRI and transverse relaxation of amyloid‐beta plaques in Alzheimer's disease. *NMR in Biomedicine,* 28**,** 297-305.

MUELLER, S. G., SCHUFF, N., YAFFE, K., MADISON, C., MILLER, B., WEINER, M. W., MUELLER, S. G., SCHUFF, N., YAFFE, K., MADISON, C., MILLER, B. & WEINER, M. W. 2010. Hippocampal atrophy patterns in mild cognitive impairment and Alzheimer's disease. *Human Brain Mapping,* 31**,** 1339-47.

NASREDDINE, Z. S., PHILLIPS, N. A., BÉDIRIAN, V., CHARBONNEAU, S., WHITEHEAD, V., COLLIN, I., CUMMINGS, J. L. & CHERTKOW, H. 2005. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society,* 53**,** 695-699.

NORTON, T. J. T., PEREYRA, M., KNIGHT, M. J., MCGARRY, B. M., JOKIVARSI, K. T., GRÖHN, O. H. J. & KAUPPINEN, R. A. 2017. Stroke onset time determination using MRI relaxation times without non-ischaemic reference in a rat stroke model. *Biomedical Spectroscopy and Imaging,* 6**,** 25-35.

PERROTIN, A., DE FLORES, R., LAMBERTON, F., POISNEL, G., JOIE, R., DE LA SAYETTE, V., MÉZENGE, F., TOMADESSO, C., LANDEAU, B., DESGRANGES, B. & CHÉTELAT, G. 2015. Hippocampal Subfield Volumetry and 3D Surface Mapping in Subjective Cognitive Decline. *Journal of Alzheimer's Disease,* 48.

PLUTA, J., YUSHKEVICH, P., DAS, S. R. & WOLK, D. A. 2012. In vivo analysis of hippocampal subfield atrophy in mild cognitive impairment via semi-automatic segmentation of T2-weighted MRI. *Journal of Alzheimer's disease,* 31**,** 85-99.

REILLY, J. F., GAMES, D., RYDEL, R. E., FREEDMAN, S., SCHENK, D., YOUNG, W. G., MORRISON, J. H. & BLOOM, F. E. 2003. Amyloid deposition in the hippocampus and

entorhinal cortex: Quantitative analysis of a transgenic mouse model. *Proceedings of the National Academy of Sciences,* 100**,** 4837-4842.
RODRIGUE, K. M., DAUGHERTY, A. M., HAACKE, E. M. & RAZ, N. 2012. The role of hippocampal iron concentration and hippocampal volume in age-related differences in memory. *Cerebral Cortex,* 23**,** 1533-41.
SARAZIN, M., CHAUVIRÉ, V., GERARDIN, E., COLLIOT, O., KINKINGNÉHUN, S., DE SOUZA, L., HUGONOT-DIENER, L., GARNERO, L., LEHÉRICY, S., CHUPIN, M. & DUBOIS, B. 2010. The amnestic syndrome of hippocampal type in Alzheimer's disease: an MRI study. *Journal of Alzheimer's disease,* 22**,** 285-294.
SELKOE, D. J. & HARDY, J. 2016. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Molecular Medicine,* 8**,** 595-608.
SMITH, M. A., ZHU, X., TABATON, M., LIU, G., JR, D. W. M., COHEN, M. L., WANG, X., SIEDLAK, S. L., DWYER, B. E., HAYASHI, T., NAKAMURA, M., NUNOMURA, A. & PERRY, G. 2010. Increased Iron and Free Radical Generation in Preclinical Alzheimer Disease and Mild Cognitive Impairment. *Journal of Alzheimer's Disease,* 19**,** 363-372.
SMITH, S. M. 2002. Fast robust automated brain extraction. *Human Brain Mapping,* 17**,** 143-155.
SØRENSEN, L., IGEL, C., PAI, A., BALAS, I., ANKER, C., LILLHOLM, M., NIELSEN, M., AGEING, F. T. A. S. D. N. I., THE AUSTRALIAN IMAGING, B. & LIFESTYLE FLAGSHIP STUDY, O. 2017. Differential diagnosis of mild cognitive impairment and Alzheimer's disease using structural MRI cortical thickness, hippocampal shape, hippocampal texture, and volumetry. *NeuroImage: Clinical,* 13**,** 470-482.
SU, L., HAYES, L., SOTERIADES, S., WILLIAMS, G., BRAIN, S. A. E., FIRBANK, M. J., LONGONI, G., ARNOLD, R. J., ROWE, J. B. & O'BRIEN, J. T. 2018. Hippocampal Stratum Radiatum, Lacunosum, and Moleculare Sparing in Mild Cognitive Impairment. *Journal of Alzheimer's disease : JAD,* 61**,** 415-424.
SYMMS, M., JÄGER, H. R., SCHMIERER, K. & YOUSRY, T. A. 2004. A review of structural magnetic resonance neuroimaging. *Journal of Neurology, Neurosurgery & Psychiatry,* 75**,** 1235-1244.
TANG, X., CAI, F., DING, D.-X., ZHANG, L.-L., CAI, X.-Y. & FANG, Q. 2018. Magnetic resonance imaging relaxation time in Alzheimer’s disease. *Brain research bulletin.*
TANG, X., HOLLAND, D., DALE, A. M., YOUNES, L., MILLER, M. I. & INITIATIVE, A. S. 2014. Shape abnormalities of subcortical and ventricular structures in mild cognitive impairment and Alzheimer's disease: detecting, quantifying, and predicting. *Human brain mapping***,** 3701-25.
TOOMBS, J. & ZETTERBERG, H. 2020. In the blood: biomarkers for amyloid pathology and neurodegeneration in Alzheimer’s disease. *Brain Communications,* 2.
VICKERS, A. J. & ALTMAN, D. G. 2001. Analysing controlled trials with baseline and follow up measurements. *BMJ,* 323**,** 1123.
WEARN, A., KNIGHT, M., MCCANN, B., MCGARRY, B., NORTON, T., JARUTYTE, L. & KAUPPINEN, R. 2017. JoVE article Matlab software. *In:* BRISTOL, U. O. (ed.).
WEARN, A. R., NURDAL, V., SAUNDERS-JENNINGS, E., KNIGHT, M. J., ISOTALUS, H. K., DILLON, S., TSIVOS, D., KAUPPINEN, R. A. & COULTHARD, E. J. 2020a. T2 heterogeneity: a novel marker of microstructural integrity associated with cognitive decline in people with mild cognitive impairment. *Alzheimer's Research & Therapy,* 12**,** 105.
WEARN, A. R., SAUNDERS-JENNINGS, E., NURDAL, V., HADLEY, E., KNIGHT, M. J., NEWSON, M., KAUPPINEN, R. A. & COULTHARD, E. J. 2020b. Accelerated long-term forgetting in healthy older adults predicts cognitive decline over 1 year. *Alzheimer's Research & Therapy,* 12**,** 119.
WESTON, P. S. J., NICHOLAS, J. M., HENLEY, S. M. D., LIANG, Y., MACPHERSON, K., DONNACHIE, E., SCHOTT, J. M., ROSSOR, M. N., CRUTCH, S. J., BUTLER, C. R.,

ZEMAN, A. Z. & FOX, N. C. 2018. Accelerated long-term forgetting in presymptomatic autosomal dominant Alzheimer's disease: a cross-sectional study. *Lancet Neurology,* 17**,** 123-132.

WOLK, D. A., DAS, S. R., MUELLER, S. G., WEINER, M. W. & YUSHKEVICH, P. A. 2017. Medial temporal lobe subregional morphometry using high resolution MRI in Alzheimer's disease. *Neurobiology of aging,* 49**,** 204-213.

XIE, L., DAS, S. R., WISSE, L. E. M., ITTYERAH, R., YUSHKEVICH, P. A., WOLK, D. A. & INITIATIVE, F. T. A. S. D. N. 2018. Early Tau Burden Correlates with Higher Rate of Atrophy in Transentorhinal Cortex. *Journal of Alzheimer's Disease,* 62**,** 85-92.

YUSHKEVICH, P. A., PLUTA, J. B., WANG, H. & XIE, L. 2015. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. *Human Brain Mapping,* 36**,** 256-87.

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 Transactions on Medical Imaging,* 20**,** 45.

ZHAO, K., DING, Y., HAN, Y., FAN, Y., ALEXANDER-BLOCH, A. F., HAN, T., JIN, D., LIU, B., LU, J., SONG, C., WANG, P., WANG, D., WANG, Q., XU, K., YANG, H., YAO, H., ZHENG, Y., YU, C., ZHOU, B., ZHANG, X., ZHOU, Y., JIANG, T., ZHANG, X., LIU, Y., INITIATIVE, A. S. D. N. & CONSORTIUM, M.-C. A. D. I. 2020. Independent and reproducible hippocampal radiomic biomarkers for multisite Alzheimer's disease: diagnosis, longitudinal progress and biological basis. *Science Bulletin,* 65**,** 1103-1113.

ZHOU, M., ZHANG, F., ZHAO, L., QIAN, J. & DONG, C. 2016. Entorhinal cortex: a good biomarker of mild cognitive impairment and mild Alzheimer's disease. *Reviews in the Neurosciences,* 27**,** 185-195.