ORIGINAL RESEARCH



A decade of changes in brain volume and cognition

Rowa Aljondi¹ · Cassandra Szoeke^{2,3} · Chris Steward¹ · Paul Yates⁴ · Patricia Desmond¹

Published online: 9 May 2018 © Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Brain atrophy can occur several decades prior to onset of cognitive impairments. However, few longitudinal studies have examined the relationship between brain volume changes and cognition over a long follow-up period in healthy elderly women. In the present study we investigate the relationship between whole brain and hippocampal atrophy rates and longitudinal changes in cognition, including verbal episodic memory and executive function, in older women. We also examine whether baseline brain volume predicts subsequent changes in cognitive performance over a 10-year period. A total of 60 individuals from the population-based Women's Healthy Ageing Project with a mean age at baseline of 59 years underwent 3T MRI. Of these, 40 women completed follow-up cognitive assessments, 23 of whom had follow-up MRI scans. Linear regression analysis was used to examine the relationship between brain atrophy and changes in verbal episodic memory and executive function over a 10-year period. The results show that baseline measurements of frontal and temporal grey matter volumes predict changes in verbal episodic memory performance, whereas hippocampal volume at baseline is associated with changes in executive function performance over a 10-year period of follow-ups. In addition, higher whole brain and hippocampal atrophy rates are correlated with a decline in verbal episodic memory. These findings indicate that in addition to atrophy rate, smaller regional grey matter volumes even 10 years prior is associated with increased rates of cognitive decline. This study suggests useful neuroimaging biomarkers for the prediction of cognitive decline in healthy elderly women.

Keywords Normal aging · Elderly women · Brain atrophy · Hippocampal atrophy · Episodic memory · Executive function

Introduction

Aging is associated with declining cognitive performance, including a decline in executive function and episodic memory (Nyberg et al. 2012; Wilson et al. 2006). These cognitive impairments are associated with synaptic loss and neural cell

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11682-018-9887-z) contains supplementary material, which is available to authorized users.

Cassandra Szoeke cszoeke@unimelb.edu.au

> Rowa Aljondi raljondi@student.unimelb.edu.au

Chris Steward csteward@unimelb.edu.au

Paul Yates Paul.YATES@austin.org.au

Patricia Desmond patricia.desmond@mh.org.au death (brain atrophy) (Morrison and Baxter 2012; Westman et al. 2013). The brain regions showing greatest volumetric reduction in grey matter in healthy elderly individuals are the frontal and temporal lobes, followed by the hippocampus in healthy adults (Bourisly et al. 2015; Lemaitre et al. 2012). Early quantification of global and regional grey matter volume

- ¹ Department of Radiology, The University of Melbourne, Royal Melbourne Hospital, Parkville, Victoria 3050, Australia
- ² Department of Medicine (Royal Melbourne Hospital), The University of Melbourne, Parkville, Victoria 3052, Australia
- ³ Institute for Health and Ageing, Australian Catholic University, Melbourne, Victoria 3000, Australia
- ⁴ Aged Care Services Department, Austin Health, Heidelberg, Victoria 3084, Australia

loss in the earlier stages of adult life could enable identification of persons at high risk of substantially accelerated cognitive decline and offer opportunities for preventive interventions.

Longitudinal structural Magnetic Resonance Imaging (MRI) has been widely used to assess how brain atrophy contributes to cognitive decline (Frisoni et al. 2010; McDonald et al. 2012; Sluimer et al. 2008). However, there is a paucity of neuroimaging studies that examine regional grey matter volumes with domain-specific cognitive decline in healthy adults. Whilst previous longitudinal studies have supported the relationship between brain atrophy and cognitive decline, they have been limited by short follow-up duration (1-5 years) (Carmichael et al. 2012; Kramer et al. 2007; Mungas et al. 2005; Rusinek et al. 2003). Others have included participants with wider age range or at older ages (> 65 years old) (Burgmans et al. 2009, Resnick et al. 2003, Sluimer et al. 2008), or individuals at risk of Alzheimer's disease (AD) with subjective memory complaints or Mild Cognitive Impairments (MCI) (Cardenas et al. 2011; Driscoll et al. 2009; Henneman et al. 2009; Jack et al. 2005; McDonald et al. 2012). These factors may limit the applicability of proposed brain volume predictors for cognitive decline in healthy adults.

The role of hippocampal atrophy in aging and cognitive decline has received significant attention in both neuroimaging and pathological research (Burgess et al. 2002; Fleischman et al. 2013; Hackert et al. 2002; Raz et al. 2005; Yavuz et al. 2007). The hippocampus is known to be affected early during the neurodegenerative process due to the early formation of the neurofibrillary tangles (Johnson et al. 2012; Rémy et al. 2005). Although the rate of hippocampal atrophy reflects the severity of cognitive decline and AD progression (Henneman et al. 2009; Yavuz et al. 2007), their relationship with domain-specific cognitive decline in healthy adults in both cross-sectional and longitudinal studies is not as clear. A large number of studies found an association between hippocampal volume and episodic memory performance in healthy adults (Cardenas et al. 2011; Carmichael et al. 2012; Hackert et al. 2002; Kramer et al. 2007; Persson et al. 2012; Tupler et al. 2007; Van Petten et al. 2004; Ystad et al. 2009). Fewer studies observed a relationship with working memory, executive function or processing speed (Fleischman et al. 2013; Oosterman et al. 2008; Papp et al. 2014). Most of these studies are limited by short follow up duration (<5 years); however, one study has assessed hippocampal atrophy and repeated neuropsychological tests over a 10-year period (den Heijer et al. 2010).

In the present study, we examine longitudinal whole brain and hippocampal volumes and subsequent changes in verbal episodic memory and executive function over a 10-year period in a cohort of healthy elderly women. We also consider baseline hippocampal volume and cortical grey matter volumes of the frontal and temporal lobes in order to test their association with longitudinal changes in cognitive performance over 10 years of follow-up assessments.

Materials and methods

Participants

All data for the present study were collected from the Women's Healthy Ageing Project (WHAP) cohort. In brief, WHAP started in 1992, when 438 Australian women (aged between 45 and 55 years) were selected by random population sampling and enrolled into a prospective longitudinal follow up study (Szoeke et al. 2013). A detailed description of the inclusion and exclusion criteria was defined in previous publication (Szoeke et al. 2016). The inclusion criteria to the initial WHAP study were women whom baseline menopausal status was able to be determined, had an intact uterus and were not taking hormonal therapy or oral contraceptives. In the present study, participants who completed brain MRI scan with cognitive assessments at baseline (2002) and follow-up (2012) time points were included in the analysis.

In 2002, a total of 60 women (mean age 59.4 ± 2.2 years) underwent brain MRI scan. Although 60 women were scanned at the baseline, the longitudinal analyses were restricted to a subsample of 40 women, who had completed cognitive follow-up measurements taken at both time points (2002 and 2012), 23 of whom had follow-up MRI scans in 2012.

The WHAP participants were selected for the present study on the basis of having completed cognitive function assessment and subsequent MRI scans across a decade of longitudinal prospective follow-ups. The study protocol for WHAP project was approved by the University of Melbourne Human Research Ethics Committee and fully complaint with the guidelines of the National Health and Medical Research Council ethical standards (HREC 931149X2 & 1,034,765). All subjects provided written informed consent before participation.

MRI acquisition

MR imaging was performed at The Royal Melbourne Hospital, using a 3T MR scanner to acquire high resolution images. In 2002, a 3T GE Signa MR imaging system was used to acquire 3D Axial T1-weighted imaging sequences with voxel dimension = $0.48 \times 0.48 \times 1.55$ mm, repetition time (TR) = 9300 ms; echo time (TE) = 1900 ms and matrix size = $512 \times 512 \times 110$. In 2012, a Siemens 3T Tim Trio MR scanner, with 12-channel phased-array head coil, was used to acquire 3D Magnetisation-Prepared Rapid Gradient Echo (MPRAGE) T1-weighted images with isotropic 1-mm voxel, TR = 1900 ms, TE = 2.98 ms and matrix size = $256 \times 256 \times 176$. Both MR sequences permit quantification of global and regional brain volumetric analysis. Scans were collected and transferred to a LINUX workstation in the brain imaging lab, at the Royal Melbourne Hospital for storage and image processing.

Imaging measures

Baseline cortical and subcortical grey matter volumes

The cortical reconstruction and volumetric segmentation of the frontal, temporal cortex and hippocampus was performed using the FreeSurfer software package, Version 5.3.0 (Laboratory for Computational Neuroimaging, Charlestown, MA, USA). The technical details of this process are described by Fischl (2012). In brief, the procedure includes skull striping, motion corrections, automated Talairach transformation and intensity normalisation. This is followed by automated topological cortical surface correction to optimally generate both the grey/white matter and the grey matter/cerebrospinal fluid interfaces (Davatzikos and Bryan 1996; Fischl et al. 1999). Surface parcellation and subcortical segmentation were completed using a predefined anatomical-labelled atlas. The reliability of FreeSufer's cortical parcellation and whole brain segmentation has been reported previously (Desikan et al. 2006; Destrieux et al. 2010; Fischl 2012; Fischl et al. 2004; Maclaren et al. 2014). Hippocampal volume measures were acquired from the segmentation, and cortical grey matter measurements of frontal and temporal lobes were obtained from cortical parcellation. The FreeSurfer output was visually checked to ensure the cortical reconstruction and subcortical segmentation of the hippocampus were accurate and without topological effects. To account for differences in head size, the output volumes were then divided by the total intracranial volume (ICV) for each subject.

Hippocampal atrophy

FreeSurfer longitudinal stream was run to estimate hippocampal atrophy from two-time points (2002 and 2012). The subcortical segmentation for the hippocampus, derived from FreeSurfer agrees well when compared to manual tracing (Cherbuin et al. 2009; Morey et al. 2009). FreeSurfer has shown higher correlation with manual tracing of the hippocampus in terms of volume overlapping (Dice's coefficient > 80%) compared to FSL/FIRST (>75%) (Wenger et al. 2014). The longitudinal assessments of hippocampal volume were performed in FreeSurfer by creating a base template from the two time point cross-sectional results. This method is known to have greatly improved the robustness and sensitivity of the overall segmentation (Morey et al. 2010; Reuter et al. 2012). The data were visually inspected for each step of FreeSufer process and corrected prior to completion of any further analysis. The hippocampal atrophy rate was calculating by subtracting the longitudinal outputs from the two-time points and expressed as percent of subjects' ICV.

Whole brain atrophy

Longitudinal whole brain volume changes were measured over a 10-year period using Structural Image Evaluation Using Normalisation of Atrophy (SIENA), part of the FMRIB-FSL software library, Version 5.0.7 (FMRIB Software Library, Oxford, UK). In brief, SIENA starts with the extraction of skull and brain images from input data at each time point (Smith et al. 2002). The two output masks derived during this process are aligned with the output of the skull to find brain/non-brain edge points between two-time points. The mean edge displacement is converted into a global estimate of the percentage of whole brain volume changes between two time intervals (Sluimer et al. 2008). This fully automated method provides a robust and accurate estimation of the percentage of brain volume changes (PBVC) and is less sensitive to inter-scanner variability of the brain images from two-time points (Cover et al. 2011; Durand-Dubief et al. 2012; Smith et al. 2002; Takao et al. 2011). Moreover, SIENA has been widely used in longitudinal MRI studies of normal aging and neurological disorders (Farzan et al. 2015; Guo et al. 2016; Mak et al. 2015; Sluimer et al. 2008; Takao et al. 2012).

Longitudinal cognitive assessments

Cognitive tests were administered to the WHAP cohort by trained psychologists. Verbal episodic memory was assessed using the delayed recall scores of the California Verbal Learning Test, second edition (CVLT-II) (Delis et al. 1987) and the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (Welsh et al. 1994). Executive function was measured by the timed Trail Making Task Part-B (TMT-B) (Clark et al. 2004a). As the Trial Making Test is a timed task with higher score indicating low performance, this variable score was transformed to a raw score by subtracting the original scores from the maximum group scores. Raw scores were then converted to z-scores based on the group mean and standard deviation (SD). Averaged composite z-scores were performed to measure verbal episodic memory. The change in verbal episodic memory and executive function over a 10-year was calculated by subtracting baseline composite z-scores from the corresponding scores 10 years later. A negative change z-score indicated a decline in cognitive performance. Details on each test's normative values and administration for this population have been reported previously (Clark et al. 2004a, b; Elkadi et al. 2006a, b; Szoeke et al. 2013). In addition, depressive symptoms were also assessed in the WHAP cohort using the Centre for Epidemiological Studies Depression (CES-D) scale. During a mean follow-up of 10 years, 3 participants developed Mild Cognitive Impairment (MCI).

Statistical analysis

All statistical analysis was performed using SPSS Version 21.0 (IPM SPSS, Chicago, IL). An independent-samples ttest was run to determine if there were differences in the baseline demographic characteristics between those women who agreed to participate in follow-up assessments compared to women who did not. Using linear regression models, initial analyses correlated the rates of whole brain and hippocampal atrophy with the z-score in changes of verbal episodic memory and executive function that occurred over a decade. Separate regression analyses were computed to model the association between baseline hippocampal, frontal and temporal grey matter volumes with changes in both cognitive functions. This test was used to predict the decline in cognitive performance (outcome variables) as a function of brain volumes (predictor variables at baseline). Linear regression models were then adjusted for age and education. These variables were chosen because they are strongly associated with cognitive functional changes and are considered to potentially confound the association between brain volume and cognition. A two-tailed *p*-value of <0.05 unadjusted for multiple comparisons was considered to be statistically significant in this study.

Results

A descriptive summary of demographic characteristics and cognitive measures for the study population at baseline is given in Table 1. The mean age of our study participants at baseline was 59.58 ± 2.25 years. They had a mean education of 13 years. Women who didn't participate in cognitive assessments at follow-up were excluded from the analysis. No differences were observed between women who included and excluded from the analysis in brain volumes, cognitive and depression scores at baseline (Table 2). However, women who participated in the follow-up analysis were better educated than those who didn't participate. Therefore, education was included in the adjusted analysis. 18.33% of our participants were APOE £4 carriers, with 11 women having at least one APOE $\varepsilon 4$ alleles at baseline. An independent sample t-test showed that there was a higher whole brain and hippocampal atrophy rates over 10 years for APOE E4 carriers compared to non-carriers, but the differences in the mean volumes were not significant (supplementary Table S1). A larger sample size of APOE $\varepsilon 4$ carriers would likely be required to further investigate this finding.

Baseline brain volumes and cognition (n = 40)

The mean of total frontal cortical volume at baseline was $11.17 \pm 0.72\%$ ICV, and total temporal cortical volume was $5.86 \pm 0.47\%$ ICV. Smaller baseline grey matter volumes of frontal and temporal lobes were associated with greater decline in verbal episodic memory a decade later (p = 0.002and p = 0.02 respectively). These findings remain significant after adjusting for age and education (Table 3). When both cortical grey matter volumes of frontal and temporal lobes were combined in the linear regression model, 21% of the verbal episodic memory scores variance was explained by the baseline frontal volume. On the other hand, we observed that smaller baseline total hippocampal volume was correlated with lower performance in executive function over a decade (p = 0.02), adjusted for age and education (Table 3). This indicates that smaller hippocampal volume predicts poorer performance in future executive function.

Longitudinal brain volume measurements and cognition (n = 23)

The mean whole brain atrophy was $0.8 \pm 0.2\%$ and hippocampal atrophy was $0.6 \pm 0.4\%$ per year. Atrophy rates were significantly correlated (r = 0.61, p = 0.002). Higher whole brain and hippocampal atrophy rates were significantly associated with decline in verbal episodic memory across a decade (p =0.001 and p = 0.03, respectively) (Table 4). When whole brain and hippocampal atrophy rates were combined in the linear regression model, 32% of the change in verbal episodic memory scores variance explained by the whole brain atrophy rate. Indeed, these results showed a slightly stronger relationship between whole brain atrophy and changes in verbal episodic memory. These findings remain significant when corrected for age and years of education for whole brain and hippocampal atrophy rates (p = 0.008 and p = 0.02), respectively (Table 4). This indicates that both higher whole brain and hippocampal atrophy rates were associated with poorer verbal memory performance over time. In contrast, no significant association was found with atrophy rates and change in executive function.

Discussion

In this prospective analysis of WHAP participants, the major findings are that baseline hippocampal volumes significantly predict changes in executive function performance, baseline grey matter volumes in frontal and temporal lobes are associated with changes in verbal episodic memory, and higher whole brain and hippocampal atrophy rates are correlated with poorer performance in verbal episodic memory across one decade.

 Table 1 Descriptive statistics for imaging cohort

Value	Baseline in 2002 $(n = 60)$ Mean $(\pm SD)$	Follow-up cognition in 2012 ($n = 40$) Mean (\pm SD)	Follow-up MRI in 2012 ($n = 23$) Mean (\pm SD)
Age, years	59.49 (2.2)	69.58 (2.2)	69.45 (2.1)
Years of education	12.83 (4.2)	13.20 (3.7)	13.21 (3.2)
% APOE ε4 carriers (n)	18.33 (11)	20.13 (8)	26.08 (6)
Cognitive test scores			
CVLT-delayed score	8.90 (3.03)	9.08 (3.05)	11.83 (4.07)
CERAD-delayed score	5.39 (2.27)	5.48 (2.26)	4.98 (2.58)
TMT-B score in second	84.11 (39.42)	84.13 (41.60)	89.56 (33.09)

Values shown are means (standard deviation) or % (n)

Abbreviation: APOE ε 4, apolipoprotein E ε 4 alleles; CVLT, California Verbal Learning Test-delayed recall score; CERAD, Consortium to Establish a Registry for Alzheimer's Disease-delayed recall score; TMT-B, Trial Making Test Part-B

Although smaller hippocampal volume at baseline is known to be sensitive in predicting episodic memory decline in older adults (Mungas et al. 2005; Tupler et al. 2007; Ystad et al. 2009), some studies have found no association with any of the cognitive domains (Cardenas et al. 2011; Carmichael et al. 2012). In our study, we observed that smaller hippocampal volume at baseline is correlated with decline in executive function over a decade of follow-up. These findings are in accordance with previous studies that have examined the relationship between hippocampal volume and executive function and processing speed, which suggests that the hippocampus may have a role beyond memory and might be involved in executive function (Oosterman et al. 2008; Papp et al. 2014). As in our study, both studies were conducted in non-demented adults' samples, but the WHAP cohort was younger at baseline (age range 56-65) and included only women. In addition, evidence from population-based studies of elderly women has shown that decline in executive function precedes decline in memory and this was associated with increased risk of late life cognitive impairments (Carlson et al. 2009; Johnson et al. 2007). Our findings imply that early measures of hippocampal volume at midlife might be a useful biomarker for predicting changes in late life executive function in elderly women.

We also observed that smaller baseline grey matter volumes in frontal and temporal lobes were associated with greater decline in verbal episodic memory a decade later. Whilst both frontal and temporal grey matter volumes are vulnerable to change with aging (Crivello et al. 2014; Driscoll et al. 2009; Lemaitre et al. 2012; Rusinek et al. 2003), their relationship with domain specific cognitive decline is inconsistent in cognitively intact subjects. In cross-sectional studies, memory performance has been found to be negatively associated with temporal and prefrontal grey matter volumes in cognitively healthy adults compared to demented individuals (Duarte et al. 2006; Van Petten 2004), whereas an opposite pattern was found in other studies examining longitudinal changes in cognition (Cardenas

Table 2 Demographic	
characteristics, cognitive	
measures and MRI volumetri	с
findings for imaging cohort a	t
baseline	

Characteristics	Included in the analysis $(n = 40)$	Excluded from the analysis $(n = 20)$	<i>p</i> -value
Age, years	59.58 (2.25)	59.45 (2.90)	0.97
Years of education	13.20 (3.78)	10.83 (2.10)	0.04
% APOE E4 carriers (n)	20.13 (8)	15.01 (3)	0.72
CVLT-delayed score	9.08 (3.05)	8.09 (2.94)	0.35
CERAD-delayed score	5.48 (2.26)	5.46 (2.24)	0.64
TMT-B score in second	84.13 (41.60)	89.10 (39.38)	0.72
CES-D score	6.02 (3.61)	6.53 (3.73)	0.76
% Frontal cortical volume /ICV	11.17 (0.72)	11.73 (0.81)	0.78
% Temporal cortical volume/ ICV	5.86 (0.47)	5.93 (0.42)	0.67
% Total hippocampal volume/ICV	0.59 (0.05)	0.58 (0.07)	0.65

Values shown are means (standard deviation) or % (n); MRI value expressed as % of Intracranial Volume (ICV) Abbreviation: APOE ε 4, apolipoprotein E ε 4 alleles; CVLT, California Verbal Learning Test-delayed recall score; CERAD, Consortium to Establish a Registry for Alzheimer's Disease-delayed recall score; TMT-B, Trial Making Test Part-B; CES-D, Centre for Epidemiological Studies Depression score (n = 40)

MRI variables (2002)	Δ Verbal episodic memory		Δ executive function	
	β	<i>p</i> -value	β	<i>p</i> -value
Univariable				
Frontal cortical volume	0.42	0.002**	0.16	0.32
Temporal cortical volume	0.38	0.02*	0.14	0.38
Total hippocampal volume	0.02	0.87	0.36	0.007*
Multivariable-adjusted ^a				
Frontal cortical volume	0.51	0.001**	0.12	0.45
Temporal cortical volume	0.37	0.02*	0.07	0.64
Total hippocampal volume	0.04	0.77	0.35	0.02*

^a adjusted for age and education

 β : standardized regression coefficient

* *p* < 0.05, ** *p* < 0.005

et al. 2011; Tisserand et al. 2004). They found that smaller grey matter volume in frontal and temporal lobes has been shown to be correlated with lower performance in both memory and executive function in elderly adults. These studies were limited by short follow-up of cognition (1–3 years), however, in our study we conducted neuropsychological tests over 10-years and observed similar findings. The role of both frontal and temporal cortices in memory is being increasingly recognised with evidence of functional connectivity of prefrontal cortex and medial temporal lobe during memory tasks (Fjell et al. 2015; Maillet and Rajah 2013). Our results thus suggest that early quantification of frontal and temporal grey matter volumes could help to identify individuals at greater risk of late life memory impairments in healthy elderly women.

A strong relationship between whole brain and hippocampal atrophy with cognitive decline has been demonstrated in MCI and demented individuals (Archer et al. 2010; den Heijer et al. 2010; Fjell et al. 2009; Henneman et al. 2009; Sluimer et al. 2008). Whilst there are several studies that examine whole brain and hippocampal atrophy with domain-specific cognitive decline in non-demented elderly subjects (Kramer et al. 2007; Mungas et al. 2005; Schmidt et al. 2005; Söderlund et al. 2004), they have been limited by short follow-up (1-6 years). Only one study assessed hippocampal atrophy and cognitive function over a 10-year period (den Heijer et al. 2010). This study examined a group of healthy elderly adults (male and female) with wider age range (60-90 years old), but also found an association between hippocampal atrophy and decline in verbal episodic memory over 10 years of follow-up. In our study, we confirm that higher whole brain and hippocampal atrophy rates are associated with greater decline in verbal episodic memory over a decade of follow-up in healthy elderly women. These findings extend the diagnostic window to 10 years before the cognitive impairment and highlight the importance of longitudinal methods for early therapeutic intervention of late life cognitive decline in healthy elderly women.

The major strength of the present study is the long-term period of our longitudinal analysis. The data were sampled from a large, prospective, longitudinal study of the WHAP cohort with follow-up assessments of cognitive function that included both verbal episodic memory performance and executive function

Table 4Linear regression forassociation between rates of brainatrophy and cognitive functionchanges over 10 years (n = 23)

MRI variables	Δ Verbal episodic memory		Δ Executive function	
	β	<i>p</i> -value	β	<i>p</i> -value
Univariable				
% whole brain atrophy	-0.63	0.001**	-0.05	0.98
% hippocampal atrophy	-0.45	0.03*	-0.08	0.71
Multivariable-adjusted ^a				
% whole brain atrophy	-0.59	0.008*	-0.16	0.87
% hippocampal atrophy	-0.61	0.02*	-0.35	0.56

^a adjusted for age and education

β: standardized regression coefficient

* p < 0.05, ** $p \le 0.005$

over a decade. The other strength of this study is that it consists of participants within a narrow age range, reducing the confounding influence of age on brain atrophy and allowing a more accurate association between brain atrophy and changes in cognitive function to be determined. In addition, automated image segmentation methods were used for longitudinal and crosssectional volumetric analyses produce similar results as found in previous studies (Apostolova et al. 2012; Morey et al. 2009; Sluimer et al. 2008; Smith et al. 2007).

Our cohort consisted only of cognitively healthy elderly women; this may limit the extent to which we can generalize our findings to a mixed population (male and female) with cognitive impairment. Although the WHAP cohort comprised a female only population, this sampling bias is important as women show greater brain atrophy and have a higher risk of developing dementia compared to men (Crivello et al. 2014; Guo et al. 2016; Hua et al. 2010). However, the small sample size included in our longitudinal study limited our ability to control for other confounding factors such as genetic variation and vascular diseases. Of the baseline cohort, only 1/3 of the participants underwent longitudinal brain MRI scans and this may have biased analyses of long term atrophic changes. Previous neuroimaging studies, with 10-year follow-up periods, had more than 50% dropout of the original cohort and included participants who developed MCI (den Heijer et al. 2010; Driscoll et al. 2009). Our study had 30% attrition with no differences between those who included and excluded from the analysis. In addition, a different scanner manufacturer in our longitudinal study may also introduce a bias in the mean volume differences. This limitation is common in longitudinal neuroimaging studies over such an extended period in terms of scanners and software upgrades (Durand-Dubief et al. 2012; Takao et al. 2011). Despite the reliability of the FreeSurfer software for the hippocampal volumes (Cherbuin et al. 2009; Morey et al. 2009), manual segmentation techniques which are the gold standard would be more accurate. Another possible limitation of our study is that cerebrovascular pathology was not directly assessed and measured. Further research is needed to clarify the relationship between cerebrovascular pathology, brain volume and cognition.

Conclusion

Structural brain volume changes identified in our crosssectional and longitudinal MRI analysis may allow early prediction of cognitive function with normal aging in elderly women. The present study suggests that midlife measures of hippocampal, cortical frontal and temporal volumes provide domain-specific contributions to the assessment of future cognitive decline. These findings may be relevant to our understanding of the neurocognitive processes with aging and provide a useful biomarker for therapeutic interventions to delay or prevent the progression of dementia. Acknowledgments We would like to acknowledge the contribution of the participants and their supporters who have contributed their time and commitment for over 20 years to the University. A full list of all researchers contributing to the project and the membership of our Scientific Advisory Board is available at http://www.medrmhwh.unimelb.edu.au/Research/WHAP.html.

Funding This study is funded by the National Health and Medical Research Council (NHMRC Grants 547500, 1032350 & 1062133), Ramaciotti Foundation, Australian Healthy Ageing Organisation, the Brain Foundation, the Alzheimer's Association (NIA320312), Australian Menopausal Society, Bayer Healthcare, Shepherd Foundation, Scobie and Claire Mackinnon Foundation, Collier Trust Fund, J.O. & J.R. Wicking Trust, Mason Foundation and the Alzheimer's Association of Australia. Inaugural funding was provided by VicHealth and the NHMRC. The Principal Investigator of WHAP (CSz) is supported by the National Health and Medical Research Council.

Compliance with ethical standards

Conflict of interest Prof. Szoeke has provided clinical consultancy and been on scientific advisory committees for the Australian Commonwealth Scientific and Industrial Research Organisation, Alzheimer's Australia, University of Melbourne and other relationships which are subject to confidentiality clauses. She has been a named Chief Investigator on investigator driven collaborative research projects in partnership with Pfizer, Merck, Bayer and GE. She has been an investigator on clinical trials with Lundbeck within the last 2 years. Dr. Desmond has been supported by the Royal Melbourne Hospital and the National Health and Medical Research Council of Australia. Other authors report no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Apostolova, L. G., Green, A. E., Babakchanian, S., Hwang, K. S., Chou, Y.-Y., Toga, A. W., & Thompson, P. M. (2012). Hippocampal atrophy and ventricular enlargement in normal aging, mild cognitive impairment and Alzheimer's disease. *Alzheimer Disease and Associated Disorders*, 26(1), 17–27.
- Archer, H., Kennedy, J., Barnes, J., Pepple, T., Boyes, R., Randlesome, K., Clegg, S., Leung, K., Ourselin, S., & Frost, C. (2010). Memory complaints and increased rates of brain atrophy: risk factors for mild cognitive impairment and Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 25(11), 1119–1126.
- Bourisly, A. K., El-Beltagi, A., Cherian, J., Gejo, G., Al-Jazzaf, A., & Ismail, M. (2015). A voxel-based morphometric magnetic resonance imaging study of the brain detects age-related gray matter volume changes in healthy subjects of 21–45 years old. *The Neuroradiology Journal*, 28(5), 450–459.
- Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, 35(4), 625–641.
- Burgmans, S., Van Boxtel, M., Smeets, F., Vuurman, E., Gronenschild, E., Verhey, F., Uylings, H., & Jolles, J. (2009). Prefrontal cortex

atrophy predicts dementia over a six-year period. *Neurobiology of Aging*, 30(9), 1413–1419.

- Cardenas, V., Chao, L., Studholme, C., Yaffe, K., Miller, B., Madison, C., Buckley, S., Mungas, D., Schuff, N., & Weiner, M. (2011). Brain atrophy associated with baseline and longitudinal measures of cognition. *Neurobiology of Aging*, 32(4), 572–580.
- Carlson, M. C., Xue, Q.-L., Zhou, J., & Fried, L. P. (2009). Executive decline and dysfunction precedes declines in memory: the women's health and aging study II. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 64(1), 110–117.
- Carmichael, O., Mungas, D., Beckett, L., Harvey, D., Farias, S. T., Reed, B., Olichney, J., Miller, J., & DeCarli, C. (2012). MRI predictors of cognitive change in a diverse and carefully characterized elderly population. *Neurobiology of Aging*, 33(1), 83–95.
- Cherbuin, N., Anstey, K. J., Réglade-Meslin, C., & Sachdev, P. S. (2009). In vivo hippocampal measurement and memory: a comparison of manual tracing and automated segmentation in a large communitybased sample. *PLoS One*, 4(4), e5265.
- Clark, M. S., Dennerstein, L., Elkadi, S., Guthrie, J. R., Bowden, S. C., & Henderson, V. W. (2004a). Normative data for tasks of executive function and working memory for Australian-born women aged 56– 67. Australian Psychologist, 39(3), 244–250.
- Clark, M. S., Dennerstein, L., Elkadi, S., Guthrie, J. R., Bowden, S. C., & Henderson, V. W. (2004b). Normative verbal and non-verbal memory test scores for Australian women aged 56–67. *Australian and New Zealand Journal of Psychiatry*, 38(7), 532–540.
- Cover, K. S., van Schijndel, R. A., van Dijk, B. W., Redolfi, A., Knol, D. L., Frisoni, G. B., Barkhof, F., Vrenken, H., & Initiative, A.s.D.N. (2011). Assessing the reproducibility of the SienaX and Siena brain atrophy measures using the ADNI back-to-back MP-RAGE MRI scans. *Psychiatry Research: Neuroimaging*, 193(3), 182–190.
- Crivello, F., Tzourio-Mazoyer, N., Tzourio, C., & Mazoyer, B. (2014). Longitudinal assessment of global and regional rate of grey matter atrophy in 1,172 healthy older adults: modulation by sex and age. *PLoS One*, 9(12), e114478.
- Davatzikos, C., & Bryan, R. N. (1996). Using a deformable surface model to obtain a shape representation of the cortex. *Medical Imaging*, *IEEE Transactions on*, 15(6), 785–795.
- Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). CVLT, California verbal learning test: adult version: manual. San Antonio: Psychological Corporation.
- den Heijer, T., van der Lijn, F., Koudstaal, P. J., Hofman, A., van der Lugt, A., Krestin, G. P., Niessen, W. J., & Breteler, M. M. (2010). A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline. *Brain*, 133(4), 1163–1172.
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., & Hyman, B. T. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3), 968–980.
- Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, 53(1), 1–15.
- Driscoll, I., Davatzikos, C., An, Y., Wu, X., Shen, D., Kraut, M., & Resnick, S. (2009). Longitudinal pattern of regional brain volume change differentiates normal aging from MCI. *Neurology*, 72(22), 1906–1913.
- Duarte, A., Hayasaka, S., Du, A., Schuff, N., Jahng, G.-H., Kramer, J., Miller, B., & Weiner, M. (2006). Volumetric correlates of memory and executive function in normal elderly, mild cognitive impairment and Alzheimer's disease. *Neuroscience Letters*, 406(1), 60–65.
- Durand-Dubief, F., Belaroussi, B., Armspach, J., Dufour, M., Roggerone, S., Vukusic, S., Hannoun, S., Sappey-Marinier, D., Confavreux, C., & Cotton, F. (2012). Reliability of longitudinal brain volume loss measurements between 2 sites in patients with multiple sclerosis:

comparison of 7 quantification techniques. American Journal of Neuroradiology, 33(10), 1918–1924.

- Elkadi, S., Clark, M. S., Dennerstein, L., Guthrie, J. R., Bowden, S. C., & Henderson, V. W. (2006a). Normative data for Australian midlife women on category fluencyand a short form of the Boston naming test. *Australian Psychologist*, *41*(1), 37–42.
- Elkadi, S., Clark, M. S., Dennerstein, L., Guthrie, J. R., Bowden, S. C., & Henderson, V. W. (2006b). Normative visuospatial performance in Australian midlife women. *Australian Psychologist*, 41(1), 43–47.
- Farzan, A., Mashohor, S., Ramli, A. R., & Mahmud, R. (2015). Boosting diagnosis accuracy of Alzheimer's disease using high dimensional recognition of longitudinal brain atrophy patterns. *Behavioural Brain Research*, 290, 124–130.
- Fischl, B. (2012). FreeSurfer. NeuroImage, 62(2), 774-781.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2), 195–207.
- Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., & Kennedy, D. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, 14(1), 11–22.
- Fjell, A. M., Walhovd, K. B., Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Holland, D., Brewer, J. B., & Dale, A. M. (2009). Oneyear brain atrophy evident in healthy aging. *The Journal of Neuroscience*, 29(48), 15223–15231.
- Fjell, A. M., Sneve, M. H., Storsve, A. B., Grydeland, H., Yendiki, A., & Walhovd, K. B. (2015). Brain events underlying episodic memory changes in aging: a longitudinal investigation of structural and functional connectivity. *Cerebral Cortex*, 26(3), 1272–1286.
- Fleischman, D. A., Yu, L., Arfanakis, K., Han, S. D., Barnes, L. L., Arvanitakis, Z., Boyle, P. A., & Bennett, D. A. (2013). Faster cognitive decline in the years prior to MR imaging is associated with smaller hippocampal volumes in cognitively healthy older persons. *Frontiers in Aging Neuroscience*, 5, 21.
- Frisoni, G. B., Fox, N. C., Jack, C. R., Scheltens, P., & Thompson, P. M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nature Reviews Neurology*, 6(2), 67–77.
- Guo, J., Isohanni, M., Miettunen, J., Jääskeläinen, E., Kiviniemi, V., Nikkinen, J., Remes, J., Huhtaniska, S., Veijola, J., & Jones, P. (2016). Brain structural changes in women and men during midlife. *Neuroscience Letters*, 615, 107–112.
- Hackert, V., den Heijer, T., Oudkerk, M., Koudstaal, P., Hofman, A., & Breteler, M. (2002). Hippocampal head size associated with verbal memory performance in nondemented elderly. *NeuroImage*, 17(3), 1365–1372.
- Henneman, W., Sluimer, J., Barnes, J., Van Der Flier, W., Sluimer, I., Fox, N., Scheltens, P., Vrenken, H., & Barkhof, F. (2009). Hippocampal atrophy rates in Alzheimer disease added value over whole brain volume measures. *Neurology*, 72(11), 999–1007.
- Hua, X., Hibar, D. P., Lee, S., Toga, A. W., Jack, C. R., Weiner, M. W., Thompson, P. M., & Initiative, A.s.D.N. (2010). Sex and age differences in atrophic rates: an ADNI study with n= 1368 MRI scans. *Neurobiology of Aging*, 31(8), 1463–1480.
- Jack, C., Shiung, M., Weigand, S., O'Brien, P., Gunter, J., Boeve, B., Knopman, D., Smith, G., Ivnik, R., & Tangalos, E. (2005). Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. *Neurology*, 65(8), 1227–1231.
- Johnson, J. K., Lui, L.-Y., & Yaffe, K. (2007). Executive function, more than global cognition, predicts functional decline and mortality in elderly women. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 62(10), 1134–1141.
- Johnson, K. A., Fox, N. C., Sperling, R. A., & Klunk, W. E. (2012). Brain imaging in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, 2(4), a006213.

- Lemaitre, H., Goldman, A. L., Sambataro, F., Verchinski, B. A., Meyer-Lindenberg, A., Weinberger, D. R., & Mattay, V. S. (2012). Normal age-related brain morphometric changes: nonuniformity across cortical thickness, surface area and gray matter volume? *Neurobiology* of Aging, 33(3), 617. e1–617. e9.
- Maclaren, J., Han, Z., Vos, S. B., Fischbein, N., & Bammer, R. (2014). Reliability of brain volume measurements: a test-retest dataset. *Scientific Data*, 1, 140037.
- Maillet, D., & Rajah, M. N. (2013). Association between prefrontal activity and volume change in prefrontal and medial temporal lobes in aging and dementia: a review. *Ageing Research Reviews*, 12(2), 479–489.
- Mak, E., Su, L., Williams, G. B., Watson, R., Firbank, M., Blamire, A. M., & O'Brien, J. T. (2015). Longitudinal assessment of global and regional atrophy rates in Alzheimer's disease and dementia with Lewy bodies. *NeuroImage: Clinical*, 7, 456–462.
- McDonald, C. R., Gharapetian, L., McEvoy, L. K., Fennema-Notestine, C., Hagler, D. J., Holland, D., Dale, A. M., & Initiative, A.s.D.N. (2012). Relationship between regional atrophy rates and cognitive decline in mild cognitive impairment. *Neurobiology of Aging*. 33(2), 242–253.
- Morey, R. A., Petty, C. M., Xu, Y., Hayes, J. P., Wagner, H. R., Lewis, D. V., LaBar, K. S., Styner, M., & McCarthy, G. (2009). A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *NeuroImage*, 45(3), 855–866.
- Morey, R. A., Selgrade, E. S., Wagner, H. R., Huettel, S. A., Wang, L., & McCarthy, G. (2010). Scan–rescan reliability of subcortical brain volumes derived from automated segmentation. *Human Brain Mapping*, 31(11), 1751–1762.
- Morrison, J. H., & Baxter, M. G. (2012). The ageing cortical synapse: hallmarks and implications for cognitive decline. *Nature Reviews Neuroscience*, 13(4), 240–250.
- Mungas, D., Harvey, D., Reed, B., Jagust, W., DeCarli, C., Beckett, L., Mack, W., Kramer, J., Weiner, M., & Schuff, N. (2005). Longitudinal volumetric MRI change and rate of cognitive decline. *Neurology*, 65(4), 565–571.
- Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, 16(5), 292–305.
- Oosterman, J. M., Vogels, R. L., van Harten, B., Gouw, A. A., Scheltens, P., Poggesi, A., Weinstein, H. C., & Scherder, E. J. (2008). The role of white matter hyperintensities and medial temporal lobe atrophy in age-related executive dysfunctioning. *Brain and Cognition*, 68(2), 128–133.
- Papp, K. V., Kaplan, R. F., Springate, B., Moscufo, N., Wakefield, D. B., Guttmann, C. R., & Wolfson, L. (2014). Processing speed in normal aging: effects of white matter hyperintensities and hippocampal volume loss. *Aging, Neuropsychology, and Cognition, 21*(2), 197–213.
- Persson, J., Pudas, S., Lind, J., Kauppi, K., Nilsson, L.-G., & Nyberg, L. (2012). Longitudinal structure–function correlates in elderly reveal MTL dysfunction with cognitive decline. *Cerebral Cortex*, 22(10), 2297–2304.
- Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., & Acker, J. D. (2005). Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cerebral Cortex*, 15(11), 1676–1689.
- Rémy, F., Mirrashed, F., Campbell, B., & Richter, W. (2005). Verbal episodic memory impairment in Alzheimer's disease: a combined structural and functional MRI study. *NeuroImage*, 25(1), 253–266.
- Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging

studies of older adults: a shrinking brain. *Society for Neuroscience*, 23(8), 3295–3301.

- Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Withinsubject template estimation for unbiased longitudinal image analysis. *NeuroImage*, 61(4), 1402–1418.
- Rusinek, H., De Santi, S., Frid, D., Tsui, W.-H., Tarshish, C. Y., Convit, A., & de Leon, M. J. (2003). Regional brain atrophy rate predicts future cognitive decline: 6-year longitudinal MR imaging study of normal aging 1. *Radiology*, 229(3), 691–696.
- Schmidt, R., Ropele, S., Enzinger, C., Petrovic, K., Smith, S., Schmidt, H., Matthews, P. M., & Fazekas, F. (2005). White matter lesion progression, brain atrophy, and cognitive decline: the Austrian stroke prevention study. *Annals of Neurology*, 58(4), 610–616.
- Sluimer, J. D., van der Flier, W. M., Karas, G. B., Fox, N. C., Scheltens, P., Barkhof, F., & Vrenken, H. (2008). Whole-brain atrophy rate and cognitive decline: longitudinal MR study of memory clinic patients 1. *Radiology*, 248(2), 590–598.
- Smith, S. M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P., Federico, A., & De Stefano, N. (2002). Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *NeuroImage*, *17*(1), 479–489.
- Smith, S. M., Rao, A., De Stefano, N., Jenkinson, M., Schott, J. M., Matthews, P. M., & Fox, N. C. (2007). Longitudinal and crosssectional analysis of atrophy in Alzheimer's disease: crossvalidation of BSI, SIENA and SIENAX. *NeuroImage*, 36(4), 1200–1206.
- Söderlund, H., Nyberg, L., & Nilsson, L. G. (2004). Cerebral atrophy as predictor of cognitive function in old, community-dwelling individuals. *Acta Neurologica Scandinavica*, 109(6), 398–406.
- Szoeke, C. E. I., Robertson, J. S., Rowe, C. C., Yates, P., Campbell, K., Masters, C. L., Ames, D., Dennerstein, L., & Desmond, P. (2013). The women's healthy ageing project: fertile ground for investigation of healthy participants 'at risk' for dementia. *International Review of Psychiatry*, 25(6), 726–737.
- Szoeke, C., Coulson, M., Campbell, S., & Dennerstein, L. (2016). Cohort profile: women's healthy ageing project (WHAP)-a longitudinal prospective study of Australian women since 1990. Women's Midlife Health, 2(1), 5.
- Takao, H., Hayashi, N., & Ohtomo, K. (2011). Effect of scanner in longitudinal studies of brain volume changes. *Journal of Magnetic Resonance Imaging*, 34(2), 438–444.
- Takao, H., Hayashi, N., & Ohtomo, K. (2012). A longitudinal study of brain volume changes in normal aging. *European Journal of Radiology*, 81(10), 2801–2804.
- Tisserand, D. J., Van Boxtel, M. P., Pruessner, J. C., Hofman, P., Evans, A. C., & Jolles, J. (2004). A voxel-based morphometric study to determine individual differences in gray matter density associated with age and cognitive change over time. *Cerebral Cortex*, 14(9), 966–973.
- Tupler, L. A., Krishnan, K. R. R., Greenberg, D. L., Marcovina, S. M., Payne, M. E., MacFall, J. R., Charles, H. C., & Doraiswamy, P. M. (2007). Predicting memory decline in normal elderly: genetics, MRI, and cognitive reserve. *Neurobiology of Aging*, 28(11), 1644– 1656.
- Van Petten, C. (2004). Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis. *Neuropsychologia*, 42(10), 1394– 1413.
- Van Petten, C., Plante, E., Davidson, P. S., Kuo, T. Y., Bajuscak, L., & Glisky, E. L. (2004). Memory and executive function in older adults: relationships with temporal and prefrontal gray matter volumes and white matter hyperintensities. *Neuropsychologia*, 42(10), 1313– 1335.
- Welsh, K. A., Butters, N., Mohs, R. C., Beekly, D., Edland, S., Fillenbaum, G., & Heyman, A. (1994). The consortium to establish

a registry for Alzheimer's disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology*, 44(4), 609–614.

- Wenger, E., Mårtensson, J., Noack, H., Bodammer, N. C., Kühn, S., Schaefer, S., Heinze, H. J., Düzel, E., Bäckman, L., & Lindenberger, U. (2014). Comparing manual and automatic segmentation of hippocampal volumes: reliability and validity issues in younger and older brains. *Human Brain Mapping*, 35(8), 4236–4248.
- Westman, E., Aguilar, C., Muehlboeck, J.-S., & Simmons, A. (2013). Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive impairment. *Brain Topography*, 26(1), 9–23.
- Wilson, R. S., Li, Y., Bienias, L., & Bennett, D. A. (2006). Cognitive decline in old age: separating retest effects from the effects of growing older. *Psychology and Aging*, 21(4), 774–789.
- Yavuz, B. B., Ariogul, S., Cankurtaran, M., Oguz, K. K., Halil, M., Dagli, N., & Cankurtaran, E. S. (2007). Hippocampal atrophy correlates with the severity of cognitive decline. *International Psychogeriatrics*, 19(04), 767–777.
- Ystad, M. A., Lundervold, A. J., Wehling, E., Espeseth, T., Rootwelt, H., Westlye, L. T., Andersson, M., Adolfsdottir, S., Geitung, J. T., & Fjell, A. M. (2009). Hippocampal volumes are important predictors for memory function in elderly women. *BioMed Central Ltd*, 9, 17.