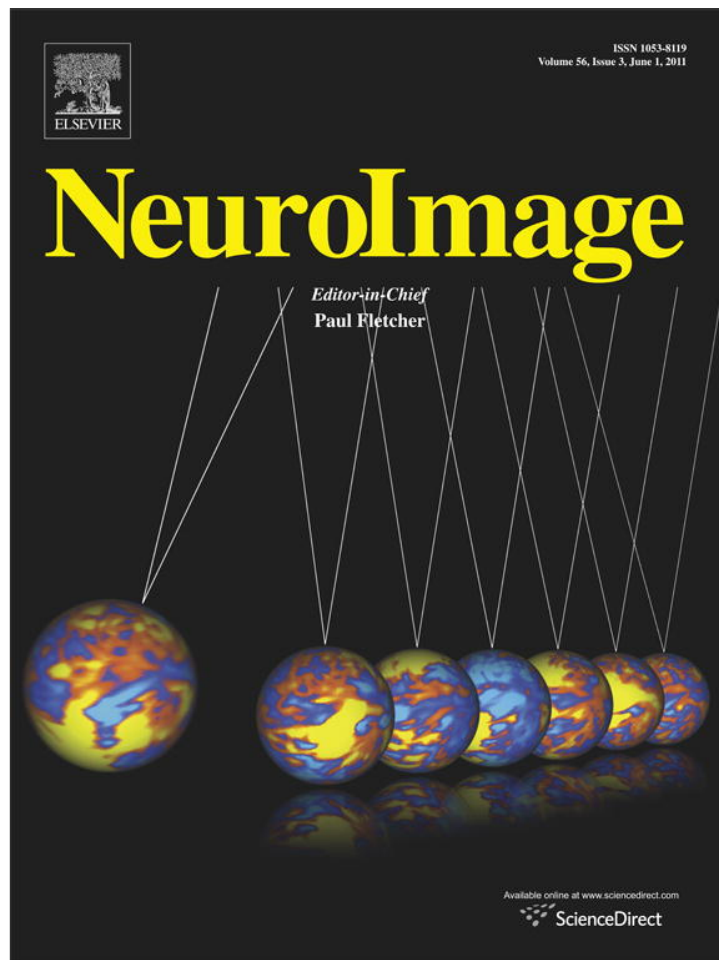


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

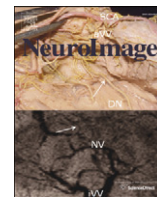
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

NeuroImage

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

Sex differences in grey matter atrophy patterns among AD and aMCI patients: Results from ADNI

Martha Skup^a, Hongtu Zhu^{b,c,*}, Yaping Wang^{b,c,d}, Kelly S. Giovanello^{c,e}, Ja-an Lin^b, Dinggang Shen^{c,f}, Feng Shi^{c,f}, Wei Gao^g, Weili Lin^{c,f}, Yong Fan^h, Heping Zhang^a and The Alzheimer's Disease Neuroimaging Initiative

^a Biostatistics Division, Yale University School of Public Health, P.O. Box 208034, 60 College Street, New Haven, CT 06520, USA

^b Department of Biostatistics, School of Public Health, The University of North Carolina at Chapel Hill, CB #7420, Chapel Hill, NC 27599, USA

^c Biomedical Research Imaging Center, The University of North Carolina at Chapel Hill, CB #7515, Chapel Hill, NC 27599, USA

^d Department of Automation, Northwestern Polytechnical University, 127 Youyi Xilu, Xi'an, Shaanxi Province 710072, PR China

^e Department of Psychology, The University of North Carolina at Chapel Hill, CB #3270, Chapel Hill, NC 27599, USA

^f Department of Radiology, The University of North Carolina at Chapel Hill, CB #7515, Chapel Hill, NC 27599, USA

^g Department of Biomedical Engineering, The University of North Carolina at Chapel Hill, CB #7515, Chapel Hill, NC 27599, USA

^h Institute of Automation, The Chinese Academy of Sciences, Beijing 100080, PR China

ARTICLE INFO

Article history:

Received 30 November 2010

Revised 14 February 2011

Accepted 20 February 2011

Available online 26 February 2011

Keywords:

Alzheimer's disease

Mild cognitive impairment

Sex differences

Longitudinal MRI

GEE

ABSTRACT

We used longitudinal magnetic resonance imaging (MRI) data to determine whether there are any gender differences in grey matter atrophy patterns over time in 197 individuals with probable Alzheimer's disease (AD) and 266 with amnesic mild cognitive impairment (aMCI), compared with 224 healthy controls participating in the Alzheimer's Disease Neuroimaging Initiative (ADNI). While previous research has differentiated probable AD and aMCI groups from controls in brain atrophy, it is unclear whether and how sex plays a role in patterns of change over time. Using regional volumetric maps, we fit longitudinal models to the grey matter data collected at repeated occasions, seeking differences in patterns of volume change over time by sex and diagnostic group in a voxel-wise analysis. Additionally, using a region-of-interest approach, we fit longitudinal models to the global volumetric data of predetermined brain regions to determine whether this more conventional approach is sufficient for determining sex and group differences in atrophy. Our longitudinal analyses revealed that, of the various grey matter regions investigated, males and females in the AD group and the aMCI group showed different patterns of decline over time compared to controls in the bilateral precuneus, bilateral caudate nucleus, right entorhinal gyrus, bilateral thalamus, bilateral middle temporal gyrus, left insula, and right amygdala. As one of the first investigations to model more than two time points of structural MRI data over time, our findings add insight into how AD and aMCI males and females differ from controls and from each other over time.

© 2011 Elsevier Inc. All rights reserved.

Introduction

As the most common form of dementia (Kukull et al., 2002), Alzheimer's disease (AD) is a neurodegenerative brain disorder that

Abbreviations: AD, Alzheimer's disease; aMCI, amnesic mild cognitive impairment; ADNI, Alzheimer's Disease Neuroimaging Initiative; MRI, magnetic resonance imaging; MAGEE, multiscale adaptive generalized estimating equation; GEE, generalized estimating equation.

* Corresponding author at: Department of Biostatistics, School of Public Health, University of North Carolina at Chapel Hill, CB #7515, Chapel Hill, NC 27599-7420, USA. Fax: +1 919 966 3804.

E-mail addresses: martha.skup@yale.edu (M. Skup), hzhu@bios.unc.edu (H. Zhu), ypwang@email.unc.edu (Y. Wang), kgio@unc.edu (K.S. Giovanello), jlin@bios.unc.edu (J. Lin), dgshen@med.unc.edu (D. Shen), fengshi@med.unc.edu (F. Shi), wgaog@email.unc.edu (W. Gao), weili.lin@med.unc.edu (W. Lin), yfan@nlpr.ia.ac.cn (Y. Fan), heping.zhang@yale.edu (H. Zhang).

leads to progressive loss in memory and cognition, and eventual death (Selkoe, 2001). A definitive diagnosis of AD can only be made postmortem after examining brain tissue for the presence of neuritic β -amyloid plaques and neurofibrillary tangles in certain brain regions (Fan et al., 2008). However, a non-invasive neuroimaging technique of magnetic resonance imaging (MRI), which provides detailed images of brain structures *in vivo*, has recently been shown to play an important role in AD diagnosis and research. Through structural MRI, distinctive and reliable biomarkers of the disease are now becoming incorporated into new diagnostic research criteria for AD (Dubois et al., 2007).

Specifically, volumetric measurements of certain brain structures collected longitudinally via a series of MRI scans provide an objective and quantitative method to examine normal aging and brain disease progression by indexing brain structural integrity and indirectly reflecting underlying neuronal health of the brain over time (Ho et al., 2010). MRI studies of normal aging brains have demonstrated significant

age-associated decreases in global and regional brain volumes including in the temporal lobe (Mueller et al., 1998; Raz et al., 1997; Schill et al., 2003; Sullivan et al., 1995), frontal lobe (Coffey et al., 1992; Raz et al., 1997), amygdala (Allen et al., 2005; Walhovd et al., 2009), thalamus (Sullivan et al., 2004; Walhovd et al., 2009), and hippocampus (Coffey et al., 1992; Schill et al., 2003; Walhovd et al., 2009), as well as increases in CSF space (Resnick et al., 2000) and the ventricles (Schill et al., 2003; Sullivan et al., 1995; Walhovd et al., 2009). While structures in the brain change in volume with normal aging, research has shown that this is greatly accelerated in AD as well as amnesic mild cognitive impairment (aMCI) (a transitional stage between normal aging and Alzheimer's disease). MRI measurements of volume loss have been used to assess atrophy in AD (Freeborough et al., 1997; O'Brien et al., 2001) and to differentiate both single-domain and multi-domain aMCI and AD from control (Jack et al., 2004; Jack et al., 2005; Sluimer et al., 2010). Efforts are ongoing to ascertain specific factors that influence differences in atrophy. Since individuals with AD and aMCI are a heterogeneous group, identification of such factors could lead to further stratifying groups at risk of decline. For instance, if AD and aMCI affect men and women in distinct ways, differentiating the sexes in research and clinical assessments may lead to more informative results and treatments.

Sex differences in normal aging and in those with dementia

Research has examined the effects of sex on aging in both normal and pathological populations. In healthy controls, studies on sexual dimorphism report that females have a higher life expectancy than their male counterparts (Nathanson, 1984; Perrig-Chiello and Hutchison, 2010). In this same population, sex differences in cognitive aging have proved inconsistent, with some investigations showing greater age-related cognitive decline in males than in females in both cross-sectional and longitudinal studies (e.g., Larrabee and Crook, 1993; Maylor et al., 2007; Meyer et al., 1999; Wiederholt et al., 1993) while others have failed to find significant sex differences in rates of cognitive aging (e.g., Barnes et al., 2003; Herlitz et al., 1997; Singer et al., 2003). Sex differences specific to AD and dementia disorders have long been suspected but are still poorly understood and controversy remains whether there are clearly recognizable differences between males and females in incidence, cognitive performance, or behavior disturbance among those afflicted (Ott and Cahn-Weiner, 2001). Previous research has consistently demonstrated that women are at higher risk of developing an incidence of AD than men (Gao et al., 1998; Vina and Lloret, 2010), which cannot be completely accounted for by women's greater life expectancy and survival advantage (Rocca et al., 1991). While some research has found that cognitive function (such as language abilities and memory performance) of women with AD compared to men is more impaired (Bai et al., 2009; Buckwalter et al., 1993; Fleisher et al., 2005; Henderson and Buckwalter, 1994; Ripich et al., 1995), other studies have reported no differences in cognition between the two gender groups (Bayles et al., 1999; Hebert et al., 2000). Further, research focusing on increased behavioral disturbance (such as aggressiveness and wandering) in AD has demonstrated qualitative differences between men and women in the manifestation of the disturbances (Drachman et al., 1992; Lyketsos et al., 1999; Ott et al., 1996; Ott et al., 2000).

Sex differences in the brain

More recently, researchers interested in sex differences among patients with AD and aMCI have turned their attention to the brain, investigating sex dimorphism in brain structure as well as in brain function by utilizing imaging technologies such as MRI, positron emission tomography (PET), and single photon emission computed tomography (SPECT) to more directly measure volumetric, physiological, and pathological changes in the brain. In healthy subjects, a growing body of literature suggests that age affects the brain differently when

comparing males and females. Sex differences in healthy individuals have been described in size, symmetry, and function of several brain structures and there is some evidence of greater age-related deterioration of the brain in one sex versus the other. For instance, in men, age-specific volume reductions are stronger than in females in whole brain volume and in the frontal and temporal lobes while women tend to show stronger reductions in the hippocampus and parietal lobe than men (Cowell et al., 1994; Nieuwenhuys et al., 2008; Sowell et al., 2007). In pathological populations with AD, imaging investigations focusing on functional brain differences between men and women have found significantly more frontal impairment in women with AD than men, more temporo-parietal impairment in men than women (Herholz et al., 2002), as well as reduced phosphorus metabolism in the frontal lobe of women with AD compared to men (Smith et al., 1995). Additionally, cross-sectional investigations of structural brain volume atrophy in men and women with AD imply that sexual dimorphism is region-specific. Research shows that AD females have smaller hippocampal volumes than males with AD, (Apostolova et al., 2006), less brain atrophy (indexed by CSF volume) in frontal, temporal, and parietal regions compared to AD men (Kidron et al., 1997), more anterior thalamic atrophy than controls (Callen et al., 2004) whereas AD males have a higher ventricle-to-brain ratio compared to women (Carmichael et al., 2007), more pronounced grey matter deficits in the anterior cingulate cortex compared to controls and AD females (Ballmaier et al., 2004), and more atrophy in the posterior cingulate compared to AD women (Callen et al., 2004). In contrast, other studies have failed to find volumetric sex differences in these populations. In a large study involving nearly 400 participants, total intracranial volume of the brain was found to be unrelated to sex in AD patients (Edland et al., 2002). In addition, brain regions not found to be differentially associated with sex in AD include the frontal lobe (Salat et al., 2001) and the medial temporal lobe (Jack et al., 1997). Studies that have investigated volumetric changes in AD and aMCI individuals differentiating sex groups *longitudinally*, over the course of several scanning sessions, are limited. One recent investigation of this nature reported that temporal lobe degeneration in AD women relative to men is faster over time (Hua et al., 2010).

Sexual dimorphism has been suggested among patients with AD and aMCI across a variety of domains, including cognition, behavior, and cross-sectional brain structure and function. However, little research has investigated longitudinal sex differences in grey matter brain volume of patients with AD and aMCI. Thus, a more thorough defining of differences between males and females over time, especially in the aMCI group, would be a useful contribution to inconclusive literature on the topic. The primary purpose of the current study was to assess whether and how atrophy in specific regions of grey matter is associated with sex for the three groups under consideration. Brain volume of various structures in the brain's grey matter were measured using serial MRI over 2–3 years in healthy control individuals, individuals with aMCI, and AD patients participating in the Alzheimer's Disease Neuroimaging Initiative (ADNI).

Materials and methods

Alzheimer's Disease Neuroimaging Initiative

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a \$60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of aMCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended

to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The principal investigator of this initiative is Michael W. Weiner, M.D., VA Medical Center and University of California–San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research—approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with aMCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years. For up-to-date information see www.adni-info.org.

Participants

We studied longitudinal brain structural changes in 687 subjects, divided into 3 groups: 224 healthy controls, 266 individuals with aMCI (both single and multi-domain), and 197 individuals with probable AD. As data collection for ADNI is ongoing, we focused on analyzing all baseline and follow-up MRI scans available at time of analysis for the subjects. On average, individuals in the AD group contributed 3.03 ($SD \pm 1.08$) scans over a 2-year period, individuals in the aMCI group contributed 3.71 ($SD \pm 1.42$) scans over a 3-year period, and healthy control individuals contributed 3.60 ($SD \pm 0.93$) scans over a 3-year period, bringing the total scans in our analysis to 2388. Our sample included only those subjects that did not change diagnostic groups throughout the duration of the study (e.g., those individuals in the aMCI group at baseline remained in the aMCI group at each follow-up scan for which we had data). Of the 832 ADNI subjects with available structural MRI data at the time of data download, 687 individuals met this “non-converter” or “stable diagnosis” requirement. We excluded 4 subjects that converted from control to aMCI, 9 subjects that converted from aMCI to control, 7 subjects that exhibited more than one conversion during the duration of their study participation, and 125 subjects that converted from aMCI to AD.

The ADNI eligibility criteria are detailed online in the ADNI protocol summary page (<http://www.adni-info.org/Scientists/ADNI-Grant/ProtocolSummary.aspx>). All subjects enrolled in this study were between the ages of 55 and 90 years of age, spoke either English or Spanish, and had an informant that was able to provide independent evaluation of the subject's functioning. Further, all subjects participated in detailed evaluations at the time of their scans assessing their clinical and neurological status as well as verifying that they still met study criteria. This included assessment of each subject's Mini Mental State Examination (MMSE) and the Clinical Dementia Rating sum-of-boxes (CDR-sob). The MMSE reflects a global measure of mental status based on evaluation of five cognitive domains (Cockrell and Folstein, 1988; Folstein et al., 1975). Scores of 24 or less are consistent with dementia, such that a score of 21–24 suggests mild dementia, 10–20 suggests moderate dementia, and less than 10 indicates severe dementia (Mungas, 1991). The CDR-sob also measures dementia severity by evaluating a patient's performance in 6 domains: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care (Hughes et al., 1982; Morris, 1993). Scores range from 0 to 18, where 10 indicates dementia. General inclusion criteria of the ADNI study were as follows: (a) healthy control subjects: MMSE scores between 24 and 30 inclusive, CDR-sob of 0, non-depressed, non-MCI, and nondemented; (b) aMCI subjects: MMSE scores between 24 and 30 inclusive, a memory complaint, objective memory loss measured by Wechsler Memory Scale Logical Memory II, absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia; and (c) probable AD: MMSE scores between 20 and 26 inclusive, a CDR-sob between 1 and 9, and met NINCDS/ADRDA criteria for probable AD (McKhann et al., 1984). Subjects were excluded if they had any serious neurological

disease other than incipient AD (such as a major depression, any history of alcohol or substance abuse), any history of brain lesions or head trauma, or psychoactive medication use.

The structural brain MRI data and corresponding clinical data from baseline and follow-up scans were downloaded on or before July 1, 2009, from the ADNI publicly available database (<http://adni.loni.ucla.edu/>). Table 1 summarizes the baseline clinical and demographic variables of the three groups.

Standard protocol patient consents

The study was conducted according to the Good Clinical Practice guidelines, the Declaration of Helsinki, US 21 CFR Part 50—Protection of Human Subjects, and Part 56—Institutional Review Boards. Subjects were willing and able to undergo test procedures, including neuroimaging and longitudinal follow-up, and written informed consent was obtained from all participants.

MRI acquisition and image corrections

The MRI data, collected across a variety of 1.5-Tesla MRI scanners with protocols individualized for each scanner, included standard T1-weighted images obtained using volumetric 3-dimensional sagittal MPRAGE or equivalent protocols with varying resolutions. As described in Jack et al. (2008), the typical protocol included repetition time (TR) = 2400 ms, inversion time (TI) = 1000 ms, flip angle = 8°, field of view (FOV) = 24 cm, with a $256 \times 256 \times 170$ acquisition matrix in the x-, y-, and z-dimensions yielding a voxel size of $1.25 \times 1.26 \times 1.2$ mm³. All original uncorrected image files are available to the general scientific community, as described at <http://www.loni.ucla.edu/ADNI>.

Image preprocessing and analysis

The image processing was carried out as follows. First, we performed AC (anterior commissure) and PC (posterior commissure) correction on all original images using MIPAV software, a free image analysis program developed by the National Institutes of Health which is available at <http://mipav.cit.nih.gov> (McAuliffe et al., 2002). In plane, we resampled the images to have dimension $256 \times 256 \times 256$. “N2” bias field correction was implemented to reduce intensity inhomogeneity in these reconstructed images (Sled et al., 2002). For each subject, we aligned the follow-up time point images to the baseline image using rigid registration for one subject in order to keep intracranial registration consistent. We performed skull-stripping on the baseline image using a hybrid of two widely used algorithms: Brain Surface Extractor (BSE) (Shattuck, et al., 2001) and Brain Extraction Tool (BET) (Smith, 2002). A number of different techniques have been developed to segment brain tissue from the skull and non-brain intracranial tissues, but it is difficult to achieve robust skull-stripping results for large-scale analyses utilizing a single method. By taking advantage of both BSE and BET, followed by manual editing, we

Table 1
Group demographics at baseline.

Group	AD subjects (n = 197)	MCI subjects (n = 266)	Control subjects (n = 224)
Age (years)	75.65 ± 7.74	74.91 ± 7.55	76.00 ± 5.00
Men/women	101/96	176/90	114/110
Education (years)	14.67 ± 3.22	15.66 ± 3.14	16.00 ± 2.88
CDR-SOB	6.18 ± 1.80	3.01 ± 0.81	1.06 ± 0.21
MMSE	23.26 ± 2.04	27.17 ± 1.78	29.12 ± 1.00

Data are given as mean ± standard deviation or n.

Key: AD, Alzheimer's disease; CDR-SOB, Clinical Dementia Rating-Sum of Boxes; MCI, milder cognitive impairment; MMSE, Mini-Mental State Exam.

were able to compensate for problems encountered with individual methods and ensure the accuracy of the skull-stripping results. Intensity inhomogeneity correction followed the skull-stripping procedure. Then, using a manually labeled cerebellum as a template, we removed the cerebellum from the images based on registration. We performed intensity inhomogeneity correction for the third time and subsequently segmented the brain into four different tissues: grey matter (GM), white matter (WM), ventricle (VN), and cerebrospinal fluid (CSF) using the FSL-FAST software (Zhang et al., 2002).

To establish the longitudinal correspondences in the individual and the intersubject correspondences between the template and the individual, we combined information across time points for each subject and registered the images to a template (Kabani et al., 1998) using a fully automatic 4-dimensional atlas warping method called 4D HAMMER (Shen and Davatzikos, 2004). We quantified the local volumetric group differences by generating RAVENS maps (Davatzikos, 1998; Davatzikos et al., 2001; Goldszal et al., 1998) for the whole brain and each of the segmented tissue type (GM, WM, VN, and CSF) using the deformation field that we obtained during registration. RAVENS methodology is based on a volume-preserving spatial transformation, which ensures that no volumetric information is lost during the process of spatial normalization since this process changes an individual's brain morphology to conform it to the morphology of a template. A physical analog is the squeezing of a rubber object, which changes the density of the rubber, to maintain the same total mass in the object (Shen and Davatzikos, 2003). Regional volumetric measurements and analyses are then performed via measurements and analyses of the resulting tissue density maps. This technique has previously been applied to a variety of longitudinal aging studies (e.g., Beresford et al., 2006; Fan et al., 2008; Resnick et al., 2000) and has been extensively validated. Lastly, we carried out automatic regional labeling: first, by labeling the template and second, by transferring the labels following the deformable registration of subject images. Labeling of the ROIs for each subject's data was done automatically and based on previously validated segmented atlas of the human brain (Kabani et al., 1998). After labeling 93 ROIs, we were able to compute volumes for each of these ROIs for each subject at each time point.

Statistical analysis

Local approach: voxel-wise analysis and cluster-level inference

Associations between brain structure, sex, and diagnostic group over time were examined by fitting a multiscale adaptive generalized estimating equation (MAGEE) model to the response measure at each voxel of the template (the RAVENS value). The MAGEE method extends the multiscale adaptive regression model (MARM) method previously described by Zhu et al. (2009) and Li et al. (in press) to the longitudinal setting, where participants are scanned repeatedly over time. It integrates a commonly used approach for analyzing longitudinal data called the generalized estimating equation (GEE) method (Liang and Zeger, 1986) with adaptive smoothing methods. GEE models account for within-subject correlation among repeated measures via the specification of a "working" correlation structure in parameter estimation. An attractive feature of GEE models is that an individual subject's response data is not necessarily assumed to be independent and, further, according to statistical theory, the estimated regression coefficients are valid even if the correlation assumption on which the analysis is based are not precisely correct.

The idea behind MAGEE essentially relies on an iteratively increasing neighborhood concept. At the beginning, a voxel-wise "naive" model is fit to the data at each voxel and naive estimates of parameters are computed based on a GEE model that assumes voxel independence. At the next step, a very small spherical "neighborhood" of every voxel is considered (which is defined by a pre-determined radius). For all voxels within the neighborhood of the voxel under consideration, weights are calculated for each neighborhood voxel

based on the estimated parameters from the previous step. The further away a neighborhood voxel is from the voxel under consideration (both in terms of physical space and in terms of similarity of the measured RAVENS value at that voxel), the smaller the weight that is assigned to the neighborhood voxel. New parameter estimates are computed at each voxel based on a weighted GEE, which integrates information of each voxel's neighborhood in estimation based on the weights. During each subsequent iteration, every neighborhood is extended by including new voxels that do not violate certain pre-set assumptions or conditions and the model estimates at each voxel are updated accordingly. New weights are computed from the estimated results obtained from the previous step and they are used to estimate the model in subsequent steps. In this way, instead of assuming independence of the voxels, this method "adaptively smoothes" the data and computes parameter estimates and test statistics by building multiple spheres around each voxels and combining all observations within the spheres with weights. MAGEE efficiently utilizes available information in the neighborhood of each voxel to increase the precision of parameter estimates and the power of test statistics in detecting subtle changes of brain structures. See Zhu et al. (2009) and Li et al. (in press) for further technical details regarding this method.

In order to validate our methodology, we first fit a simple model (hereinafter referred to as the "validation model") to the data using the above described procedure to show that our method could replicate one of the most robust findings shown in this group of subjects, namely, the between-group differences in hippocampal volume decrease over time. To do so, we treated the RAVENS measure (reflecting local volume) at each voxel as the response and included baseline intracranial volume (ICV), education, sex, age, age2, and diagnosis status (aMCI, AD, or healthy control), as well as the two-way interactions between age, age2, and diagnostic status as the model covariates. As is typical in statistical brain mapping, we applied the MAGEE procedure to each voxel within the GM RAVENS maps separately, using this validation model. Thus, we obtained estimates for the validation model's coefficients and their standard errors at each voxel, from which we calculated multivariate Wald test statistics and associated *p*-values, showing the significance of covariate effects at each voxel.

To address our scientific question of interest and identify differences in patterns of change in local brain volume over time by sex and diagnostic group, we fit a second model to the data (hereinafter referred to as the "hypothesis model"), which included baseline intracranial volume (ICV), education, sex, age, age2, and diagnosis status as well as the two-way interactions between sex, age, age2, and diagnostic status as the model covariates. Although not of primary interest, the covariates of ICV and education were chosen for theoretical reasons, based on prior hypotheses that they would have an effect, since these variables have been previously found to correlate with individual differences in brain volume (Brun et al., 2009; Luders et al., 2009). We included a quadratic age term in our model because brain atrophy has been shown to vary nonlinearly with age in elderly populations (Schuff et al., 2010). Moreover, in order to identify differences in patterns of local atrophy over time by sex and diagnostic group, we added three-way interactions (age \times sex \times diagnosis and age2 \times sex \times diagnosis) to the hypothesis model. We again applied the MAGEE procedure to each voxel within the GM RAVENS maps separately, obtaining estimates for the hypothesis model's coefficients and their standard errors at each voxel, from which we calculated multivariate Wald test statistics and associated *p*-values.

Our primary scientific interest was to localize distinct patterns of atrophy in brain structure for sex and diagnostic groups over time. Additionally, to validate our methodology, we aimed to replicate findings of increased hippocampal atrophy over time in AD and aMCI patients compared to controls. With these two goals in mind, we limited our initial focus to only those Wald statistic maps which tested

for the presence of an age by diagnostic group interaction in the case of the validation model and a sex by age by diagnostic group interaction in the case of the hypothesis model. We generated *p*-value maps by performing two Wald tests at each voxel, which tested the hypotheses (1) that the time by diagnosis interaction at that voxel was equal to zero and (2) that the time by sex by diagnosis interaction at that voxel was equal to zero, for the validation model and hypothesis model, respectively. Then, we thresholded the maps at *p*-value < 0.01, uncorrected for multiple comparisons, with an extent threshold of ≥ 20 contiguous voxels, in a cluster-level inference approach. The primary threshold was set to a significance level of 0.01 in order find a compromise between sensitivity to cluster extent and separation of maximal voxel results. Other thresholding approaches, such as false discovery rate (Benjamini and Hochberg, 1995) and random field theory (Friston et al., 1995) are also available. The anatomical localization of each cluster was established by comparing the coordinates of the cluster to the labeled template.

To determine if we could replicate the robust hippocampal finding documented in previous studies, we identified the peak voxel within clusters that spanned the hippocampus which showed evidence of a time by diagnosis interaction. We then conducted post hoc Wald tests to compare AD to controls, aMCI to controls, and AD to aMCI in hippocampal atrophy. Next, to investigate the effects of sex, *a priori*, we selected 24 regions-of-interest (ROIs) within GM that have been reported by previous research to be important in Alzheimer's disease progression (e.g., Fan et al., 2008; Fennema-Notestine et al., 2009; Ho et al., 2010; Holland et al., 2009; Jack et al., 2005). Regions included the following (note that the numbers in parenthesis indicate the number of $1 \times 1 \times 1$ mm³ voxels that made up the respective region): bilateral insula (right: 1123; left: 1216), bilateral hippocampus (right: 465; left: 572), bilateral parahippocampal gyrus (right: 390; left: 326), bilateral cingulate (right: 2118; left: 2418), bilateral precuneus

(right: 539; left: 613), bilateral middle temporal gyrus (right: 5028; left: 3340), bilateral inferior temporal gyrus (right: 1839; left: 2082), bilateral caudate nucleus (right: 633; left: 689), bilateral entorhinal gyrus (right: 256; left: 150), bilateral thalamus (right: 1218; left: 1169), bilateral uncus (right: 171; left: 238), and bilateral amygdala (right: 227; left: 139). At the peak voxel within each cluster that spanned the selected regions, we conducted five post hoc analyses, each of which utilized a subset of the original dataset. In these secondary analyses, we limited the sample to (i) only females, (ii) only males, (iii) only individuals with AD, (iv) only individuals with aMCI, or (v) only healthy controls, respectively. These analyses allowed us to identify the driving force behind the significant three-way interactions found in the primary analysis. We again applied the MAGEE procedure in an identical fashion to the one previously described, this time using a reduced model and reduced sample to obtain Wald statistics and *p*-value maps for each effect of interest in the reduced model. No post hoc analyses were corrected for multiple comparisons.

Global approach: region-of-interest analysis

In addition to an analysis at a localized voxel level, we pursued an analysis at a more spatially coarse global level, targeting an entire anatomical brain region corresponding to the underlying anatomical parcellation of the brain. Of the 93 regions for which we had ROI volume data over time, we again pre-selected 24 GM ROIs for the analysis. To identify differences in patterns of ROI volumetric change over time by sex and diagnostic group, we treated the volume of each ROI as the dependent variable and fit a generalized estimating equation (GEE) model (the hypothesis model) with a AR(1) working correlation structure. Just as we did in the MAGEE voxel-wise analysis, here, we included in the model the following covariates: baseline ICV, education, sex, age, age2, and diagnosis status, two-way interactions

Table 2
The estimated values (\pm standard errors) for the main coefficients of interest in the overall Wald test for the local RAVENS analysis and the global analysis. The top two sections correspond to the validation model and the bottom two sections correspond to the hypothesis model.

Validation model				
Region of interest	Coefficients			
	AD sex time interaction	AD sex time2 interaction	MCI time interaction	MCI time2 interaction
<i>Ravens local analysis</i>				
L. hippocampus	-1.529 \pm 0.497	0.097 \pm 0.066	0.138 \pm 0.464	-0.059 \pm 0.062
R. hippocampus	-1.990 \pm 0.503	0.147 \pm 0.066	-0.325 \pm 0.499	-0.003 \pm 0.066
<i>Global analysis</i>				
L. hippocampus	-3902.87 \pm 1699.56	243.38 \pm 108.36	-4290.51 \pm 1435.06	275.30 \pm 90.64
R. hippocampus	-3685.75 \pm 1658.76	220.46 \pm 104.96	-3518.63 \pm 1540.39	216.86 \pm 97.03
Hypothesis model				
Region of interest	Coefficients			
	AD sex time interaction	AD sex time2 interaction	MCI sex time interaction	MCI sex time2 interaction
<i>Ravens local analysis</i>				
L. thalamus	-1.280 \pm 0.305	0.178 \pm 0.041	-0.207 \pm 0.345	0.016 \pm 0.046
R. precuneus	-0.819 \pm 0.389	0.141 \pm 0.053	-0.581 \pm 0.404	0.125 \pm 0.053
R. entorhinal gyrus	-0.414 \pm 0.431	0.074 \pm 0.059	-0.474 \pm 0.374	0.117 \pm 0.051
R. middle temporal gyrus	-1.064 \pm 0.393	0.098 \pm 0.053	-0.100 \pm 0.393	-0.026 \pm 0.053
R. amygdala	-0.109 \pm 0.280	-0.004 \pm 0.039	-0.811 \pm 0.222	0.101 \pm 0.030
R. thalamus	-1.028 \pm 0.276	0.141 \pm 0.037	0.021 \pm 0.259	-0.022 \pm 0.036
<i>Global analysis</i>				
L. caudate	724.33 \pm 2404.44	-34.33 \pm 160.25	4952.23 \pm 2217.38	-331.88 \pm 148.45
L. insula	6089.21 \pm 7349.29	-379.10 \pm 469.94	14646.27 \pm 7045.33	-951.89 \pm 448.32
L. middle temporal gyrus	9615.32 \pm 7826.28	-624.46 \pm 505.88	18283.14 \pm 6061.78	-1158.60 \pm 392.07
L. precuneus	756.05 \pm 1524.23	-41.49 \pm 98.52	3438.12 \pm 1327.09	-224.64 \pm 85.86
L. thalamus	-3496.07 \pm 2883.07	231.05 \pm 183.98	-3705.50 \pm 2813.03	229.24 \pm 179.18
R. caudate	3156.72 \pm 2383.59	-192.09 \pm 158.35	6721.47 \pm 2045.22	-442.12 \pm 135.59
R. middle temporal gyrus	1455.96 \pm 7772.70	-182.225 \pm 513.38	11567.24 \pm 6479.95	-752.228 \pm 433.85

between sex, age, age2, and diagnosis, as well as two three-way interactions (age × sex × diagnosis and age2 × sex × diagnosis). Post hoc analyses were conducted on subsets of the dataset (restricted samples of (i) only females, (ii) only males, (iii) only individuals with AD, (iv) only individuals with aMCI, or (v) only healthy controls) in order to better understand significant results of the primary analysis. Additionally, to validate our method, as we did in the local analysis, we fit a simpler GEE model to the global hippocampus volume data, which included the following covariates: baseline ICV, education, sex, age, age2, diagnosis status, and two-way interactions between sex, age, age2, and diagnosis. We used this validation model to identify differences in patterns of ROI volumetric change over time by diagnostic group.

Results

Validation local and global analysis: group comparisons in hippocampal atrophy

We tested for a diagnosis effect in local atrophy patterns over time within clusters that spanned the hippocampus using a voxel-wise analysis and cluster-level inference. To do so, we used Wald-statistic maps generated by fitting MAGEE validation models to RAVENS grey matter (GM) maps. Specifically, Wald tests at the voxel with the maximum Wald statistic revealed differences in the right and left hippocampus atrophy between AD and controls ($W = 228.4565$

($df = 2$), p -value < 0.0001 and $W = 219.0538$ ($df = 2$), p -value < 0.0001, respectively), aMCI and controls ($W = 30.5029$ ($df = 2$), p -value < 0.0001 and $W = 29.8333$ ($df = 2$), p -value < 0.0001, respectively), and AD and aMCI ($W = 76.8145$ ($df = 2$), p -value < 0.0001 and $W = 78.8579$ ($df = 2$), p -value < 0.0001, respectively). Similarly, we tested for a diagnosis effect in global atrophy patterns over time in the hippocampus by fitting GEE models to global hippocampal ROI volume data. Here, Wald tests revealed differences in the right and left hippocampus ROI volume over time between AD and controls ($W = 10.15$ ($df = 2$), p -value = 0.0062 and $W = 6.17$ ($df = 2$), p -value = 0.0456, respectively) as well as aMCI and controls ($W = 6.09$ ($df = 2$), p -value = 0.0476 and $W = 9.90$ ($df = 2$), p -value = 0.0071, respectively). Our findings showed that AD and aMCI patients had decreased hippocampal volume over time compared to controls, thereby replicating the robust finding. The parameter estimates for the covariates central in the construction of these Wald statistics are shown in Table 2. Fig. 1 shows mean profiles of the fitted local and global volume measures for the right and left hippocampus determined by the Wald statistic to be significant.

To give an indication regarding the fit of the validation model (above) compared to the fit of the hypothesis model (the focus of the subsequent results section), the Quasi-likelihood Independence Criterion (QIC) statistic (Pan, 2001) was utilized. The QIC is analogous to the Akaike's Information Criterion (AIC) (Akaike, 1973) in that it can be used to determine whether inclusion of covariates significantly improves the fit of a GEE model. Our assessment revealed that for the left hippocampus, QIC = 2492.6497 for the validation model and QIC = 2450.2359 for the hypothesis model. In the right hippocampus,

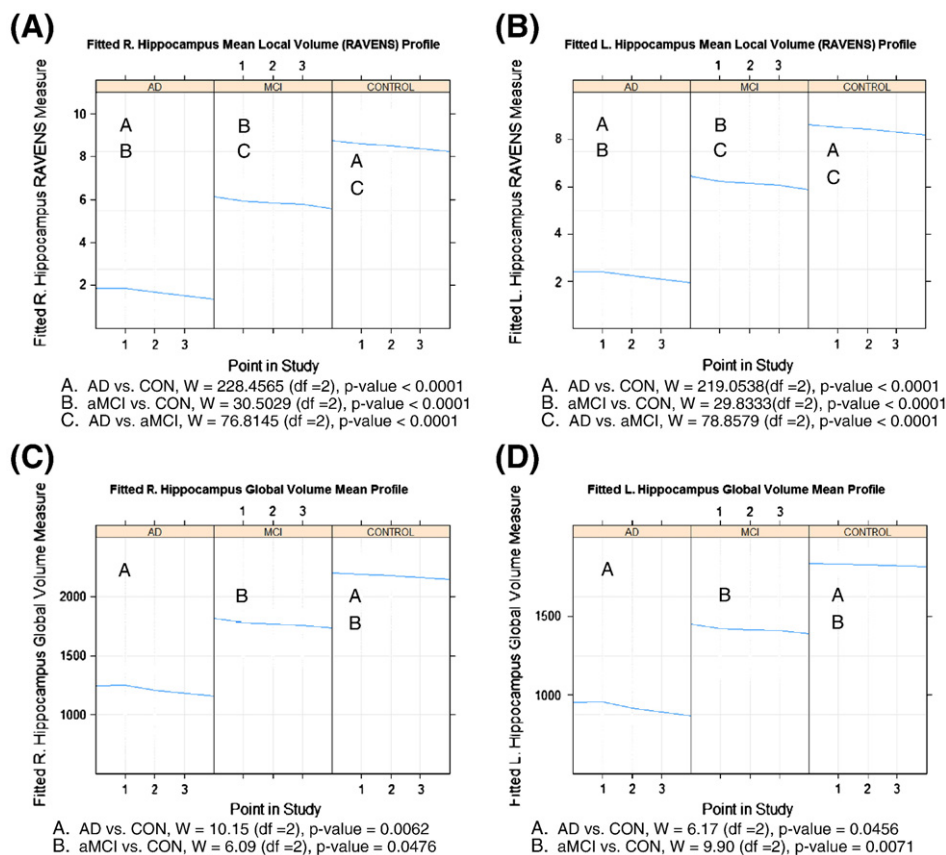


Fig. 1. Fitted mean brain volumes, corrected for ICV, for the right and left hippocampus over time, by diagnostic group, from the global and local analyses. To correct for ICV, fitted values were corrected as follows: corrected value = fitted value - β ICV (where β is the estimated coefficient from the estimated model corresponding to ICV). In the global analysis plots, the fitted value represents the estimated volume of that region based on the global analysis GEE model. In the RAVENS analysis plots, the fitted value represents the estimated RAVENS measure of the voxel based on the local analysis MAGEE model. The x-axis represents time and the y-axis represents the corrected fitted mean brain volumes as defined by the equation. Significant pairwise differences between groups are noted.

QIC = 2446.7483 for the validation model and QIC = 2492.2944 for the hypothesis model.

Primary local analysis: group comparisons via voxel-wise analysis of RAVENS GM maps

Using Wald-statistic maps generated by fitting MAGEE hypothesis models to RAVENS grey matter (GM) maps, we tested for sex differences across the three diagnostic groups in local atrophy patterns over time within clusters that spanned 24 GM ROIs in a voxel-wise analysis with cluster-level inference. Our study yielded sex and gender differences over time in local volume change at clusters across 6 of these regions. Specifically, Wald tests at the voxel with the maximum Wald statistic of each cluster revealed significant group by sex by age interactions in the right precuneus ($W = 28.6242$ ($df = 4$), p -value < 0.0001), right entorhinal gyrus ($W = 28.6641$ ($df = 4$), p -value < 0.0001), right and left thalamus ($W = 24.1266$ ($df = 4$), p -value = 0.0001 and $W = 22.6539$ ($df = 4$), p -value = 0.0002, respectively), right middle temporal gyrus ($W = 22.2619$ ($df = 4$), p -value = 0.0002), and right amygdala ($W = 19.5899$ ($df = 4$), p -value = 0.0007). Table 2 depicts the parameter estimates for the covariates most important for constructing these Wald statistics. Fig. 2 shows the uncorrected $-\log_{10}(p)$ -value maps corresponding to the three-way interactions in the relevant regions.

Secondary analyses were conducted to determine how males and females differed in their local atrophy patterns for the regions identified by the primary analysis. We again fit MAGEE models to RAVENS GM maps, but this time using subsets of the original dataset. Five such post hoc analyses were carried utilizing data from (i) only the AD group, (ii) only the aMCI group, (iii) only the healthy control group, (iv) only the males, and (v) only the females.

Secondary local analyses: sex groups

An analysis of only the male subjects revealed that males with AD showed significantly different patterns of local volumetric decline than controls in 4 of the 6 regions determined to be important by the primary analysis. Specifically, Wald tests revealed differences in the following regions: the right precuneus ($W = 8.3309$ ($df = 2$), p -value = 0.0164), right middle temporal gyrus ($W = 42.3039$ ($df = 2$), p -value < 0.0001), right and left thalamus ($W = 7.2863$ ($df = 2$), p -value = 0.0273 and $W = 7.8784$ ($df = 2$), p -value = 0.0204, respectively), and right amygdala ($W = 111.8166$ ($df = 2$), p -value < 0.0001). Differences in local volumetric decline in aMCI males compared to controls were shown in the right entorhinal gyrus ($W = 6.1834$ ($df = 2$), p -value = 0.0469), right middle temporal gyrus ($W = 10.447$ ($df = 2$), p -value = 0.0058), right amygdala ($W = 32.9219$ ($df = 2$), p -value < 0.0001), and right precuneus ($W = 7.6006$ ($df = 2$), p -value = 0.0234). This same analysis demonstrated that AD males displayed significantly different patterns of local volume decline over time than aMCI males in the right and left thalamus ($W = 9.7586$ ($df = 2$), p -value = 0.0082 and $W = 14.9028$ ($df = 2$), p -value = 0.0007, respectively), right middle temporal gyrus ($W = 17.3434$ ($df = 2$), p -value = 0.0002), and right amygdala ($W = 34.3651$ ($df = 2$), p -value < 0.0001). A summary of these results is presented in Table 3. Mean profiles of fitted local volume measures for each region determined by the Wald statistic to be significant in this male-only analysis are shown in Fig. 3.

A female-only investigation revealed differences between female individuals with AD as compared to female controls in 3 of the 6 regions. These regions included the right precuneus ($W = 17.9977$ ($df = 2$), p -value = 0.0002), right thalamus ($W = 6.7376$ ($df = 2$), p -value = 0.0362), and right amygdala ($W = 67.6891$ ($df = 2$), p -value < 0.0001). This female-only analysis also demonstrated differences between aMCI females and control females in the right precuneus ($W = 23.9969$ ($df = 2$), p -value < 0.0001), right entorhinal gyrus ($W = 25.5225$ ($df = 2$), p -value $<$

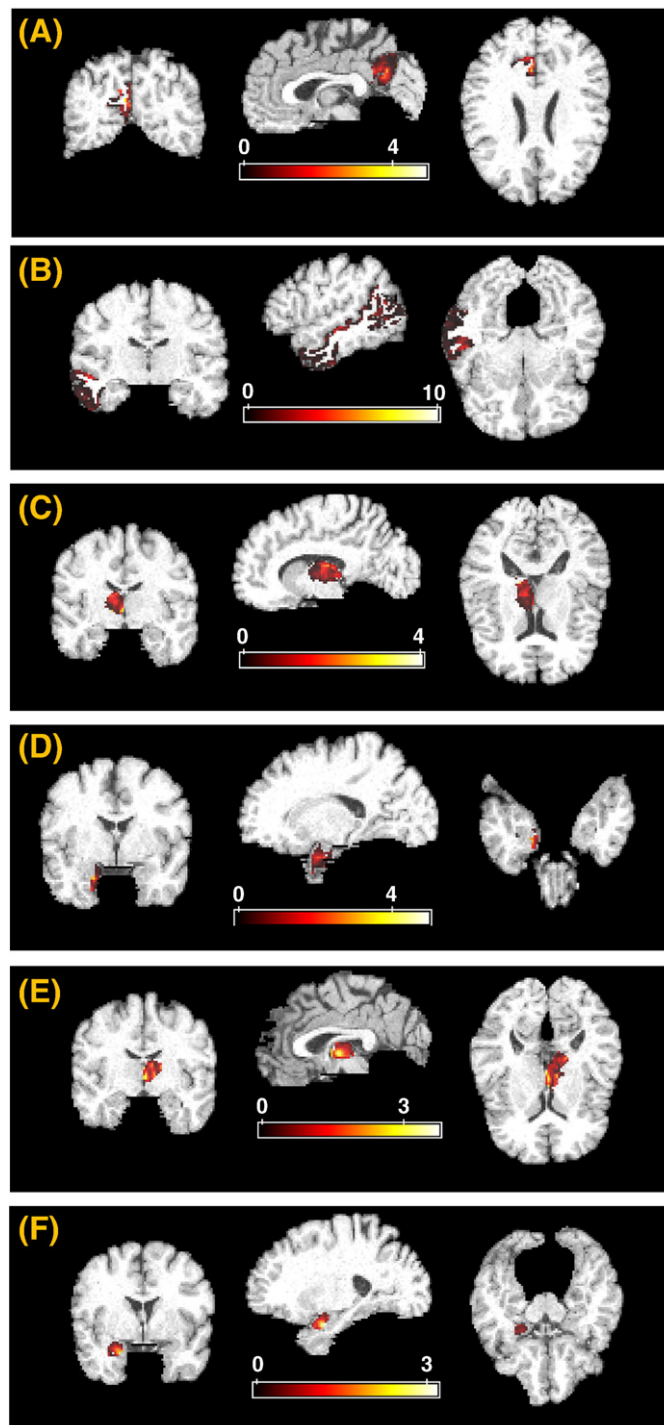


Fig. 2. Voxel-based analysis of RAVENS maps, showing $-\log_{10}(p)$ values from the primary analysis for various regions of interest displayed on the original intensity template. (A) Right precuneus (MNI: 2, -58, 24); (B) right middle temporal gyrus (MNI: 54, -22, -8); (C) right thalamus (MNI: 16, -24, 8); (D) right entorhinal cortex (20, -2, -26); (E) left thalamus (MNI: -4, -8, 2); (F) right amygdala (MNI: 22, -8, -14). The color maps indicate the scale for the $-\log_{10}(p)$ values.

0.0001), and right amygdala ($W = 18.0029$ ($df = 2$), p -value = 0.0002). AD females and aMCI females differed in local volumetric decline in the right and left thalamus ($W = 12.5327$ ($df = 2$), p -value = 0.0022 and $W = 7.3562$ ($df = 2$), p -value = 0.0268, respectively), right entorhinal gyrus ($W = 9.8476$ ($df = 2$), p -value = 0.008), right amygdala ($W = 18.8435$ ($df = 2$), p -value = 0.0001), and right middle temporal

Table 3

Results from two secondary analyses, using a male-only subset of the data (displayed on the left) and a female-only subset of the data (displayed on the right). Only significant Wald statistics, which correspond to a test of diagnostic effect over time, are shown for relevant regions of interest. “Local” refers to the local RAVENS map MAGEE analysis and “global” refers to the global ROI volume GEE analysis.

Males only					Females only				
Region of interest	Wald stat (<i>df</i> =2)	<i>p</i> -value	Analysis	Pattern	Region of interest	Wald stat (<i>df</i> =2)	<i>p</i> -value	Analysis	Pattern
<i>AD vs. healthy control</i>					<i>AD vs. healthy control</i>				
L. precuneus	5.37	0.0681	Global	AD<Con	L. thalamus	8.48	0.0144	Global	AD<Con
L. thalamus	7.8784	0.0204	Local	AD<Con	R. amygdala	67.6891	<0.0001	Local	AD<Con
L. thalamus	5.16	0.0759	Global	AD<Con	R. caudate nucleus	6.14	0.0464	Global	AD<Con
R. amygdala	111.8166	<0.0001	Local	AD<Con	R. precuneus	17.9977	0.0002	Local	AD<Con
R. middle temporal gyrus	42.3039	<0.0001	Local	AD<Con	R. middle temporal gyrus	5.63	0.0598	Global	AD<Con
R. precuneus	8.3309	0.0164	Local	AD<Con	R. thalamus	6.7376	0.0362	Local	AD<Con
R. thalamus	7.2863	0.0273	Local	AD<Con					
<i>aMCI vs. healthy control</i>					<i>aMCI vs. healthy control</i>				
L. middle temporal gyrus	5.26	0.0721	Global	aMCI<Con	L. middle temporal gyrus	6.46	0.0396	Global	aMCI<Con
R. amygdala	32.9219	<0.0001	Local	aMCI<Con	R. amygdala	18.0029	0.0002	Local	aMCI<Con
R. entorhinal gyrus	6.1834	0.0469	Local	Con<aMCI	R. caudate nucleus	6.99	0.0303	Global	aMCI<Con
R. middle temporal gyrus	10.447	0.0058	Local	aMCI<Con	R. entorhinal gyrus	25.5225	<0.0001	Local	aMCI<Con
R. precuneus	7.6006	0.0234	Local	Con<aMCI	R. middle temporal gyrus	6.41	0.0406	Global	aMCI<Con
					R. precuneus	23.9969	<0.0001	Local	aMCI<Con
<i>AD vs. aMCI</i>					<i>AD vs. aMCI</i>				
L. precuneus	10.77	0.0046	Global	AD<aMCI	L. middle temporal gyrus	9.74	0.0077	Global	AD<aMCI
L. thalamus	14.9028	0.0007	Local	AD<aMCI	L. thalamus	7.3562	0.0268	Local	AD<aMCI
R. amygdala	34.3651	<0.0001	Local	AD<aMCI	L. thalamus	7.700	0.0213	Global	AD<aMCI
R. thalamus	9.7586	0.0082	Local	AD<aMCI	R. amygdala	18.8435	<0.0001	Local	AD<aMCI
R. middle temporal gyrus	17.3434	0.0002	Local	AD<aMCI	R. entorhinal gyrus	9.8476	0.0080	Local	aMCI<AD
R. middle temporal gyrus	6.12	0.0468	Global	AD<aMCI	R. middle temporal gyrus	10.4053	0.0061	Local	AD<aMCI
					R. middle temporal gyrus	6.28	0.0432	Global	AD<aMCI
					R. thalamus	12.5327	0.0022	Local	AD<aMCI

gyrus ($W = 10.4053$ ($df = 2$), p -value = 0.0061). A summary of these results is presented in Table 3. Mean profiles of fitted local volume measures for each region determined by the Wald statistic to be significant in this female-only analysis are shown in Fig. 3.

Secondary local analyses: diagnostic groups

Three additional secondary analyses were conducted, separately for each of the diagnostic groups. An analysis of only AD individuals revealed that males and females showed varying patterns of local atrophy in the right and left thalamus ($W = 11.0063$ ($df = 2$), p -value = 0.0049 and $W = 24.0964$ ($df = 2$), p -value < 0.0001, respectively) and right middle temporal gyrus ($W = 6.596$ ($df = 2$), p -value = 0.0397). In the aMCI group, on the other hand, sex dimorphism was shown in the right and left thalamus ($W = 25.7772$ ($df = 2$), p -value < 0.0001 and $W = 25.0305$ ($df = 2$), p -value < 0.0001, respectively), right precuneus ($W = 9.5285$ ($df = 2$), p -value = 0.0094), right entorhinal gyrus ($W = 27.442$ ($df = 2$), p -value < 0.0001), and right middle temporal gyrus ($W = 8.7948$ ($df = 2$), p -value = 0.0134). Finally, female controls and male controls differed in local brain atrophy patterns in the right precuneus ($W = 20.2069$ ($df = 2$), p -value = 0.0001), right amygdala ($W = 8.8576$ ($df = 2$), p -value = 0.0133), and left thalamus ($W = 9.5174$ ($df = 2$), p -value = 0.0097) in an analysis that only included healthy control individuals. A summary of these results is presented in Table 4. Mean profiles of fitted local volume measures for each region determined by the Wald statistic to be significant in this diagnosis-only analysis are shown in Fig. 3.

Primary global analysis: group comparisons via global ROI volumetric analysis

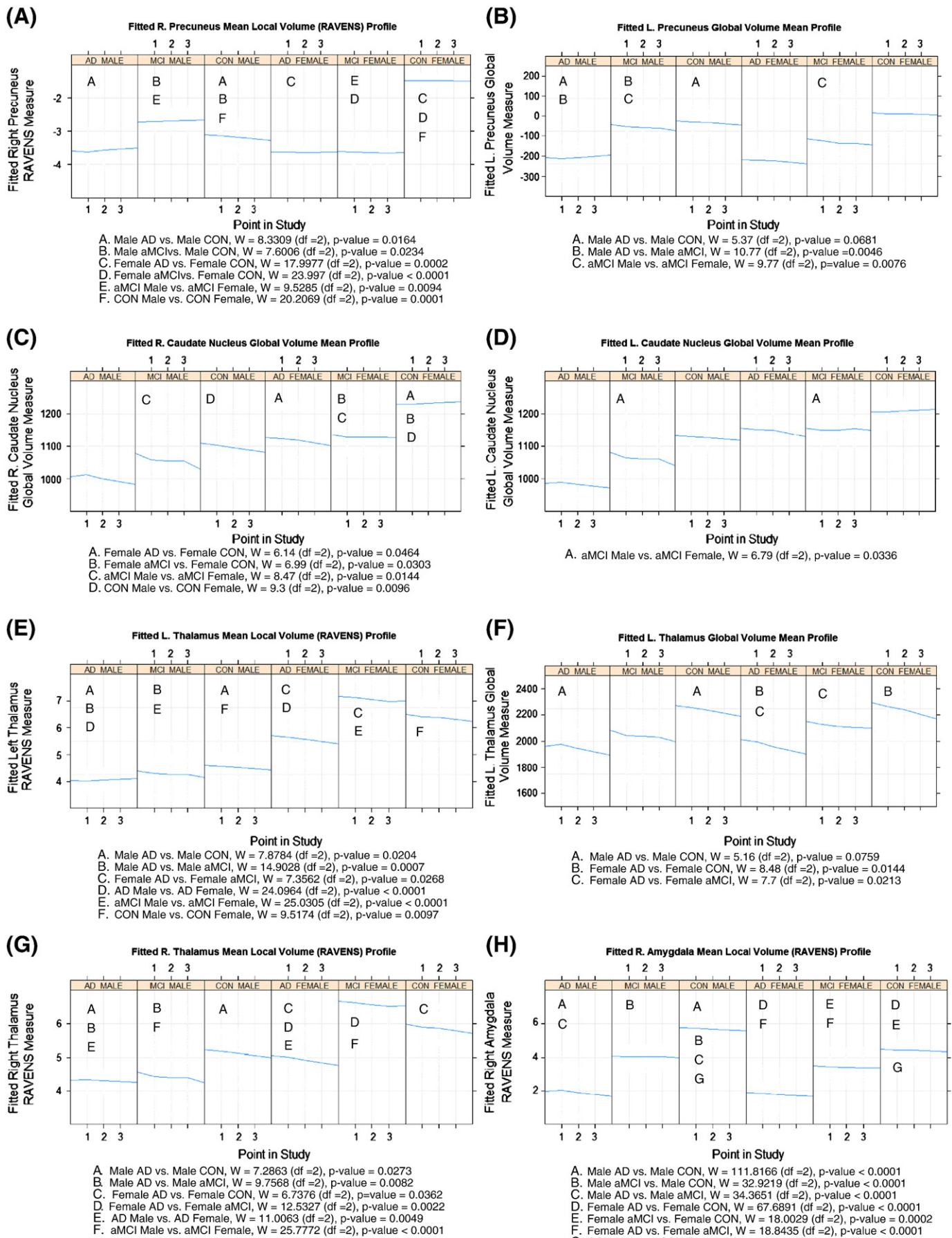
We fit GEE models to the global volume data of 24 GM ROIs and tested for sex differences across the three diagnostic groups in atrophy patterns over time in an ROI-based analysis. Wald tests revealed significant group by gender by age interactions in the left caudate

nucleus ($W = 10.39$ ($df = 4$), p -value = 0.0343), right caudate nucleus ($W = 14.00$ ($df = 4$), p -value = 0.0073), left precuneus ($W = 11.34$ ($df = 4$), p -value = 0.023), left middle temporal gyrus ($W = 12.51$ ($df = 4$), p -value = 0.014), left thalamus ($W = 9.15$ ($df = 4$), p -value = 0.0574), left insula ($W = 11.17$ ($df = 4$), p -value = 0.0248), and right middle temporal gyrus ($W = 15.85$ ($df = 4$), p -value = 0.0032). The parameter estimates for the covariates central in the construction of these Wald statistics are shown in Table 2.

Secondary analyses were conducted to determine how the gender and diagnostic groups differed in their patterns of global ROI volume decline in each of these identified regions. We again fit GEE models to the ROI volumetric data, but this time, we used subsets of the original dataset. Five such post hoc analyses were carried out utilizing data from (i) only the AD group, (ii) only the aMCI group, (iii) only the healthy control group, (iv) only the males, and (v) only the females.

Secondary global analyses: sex groups

An analysis of only females revealed that female subjects with aMCI demonstrated significantly different patterns and acceleration of decline than control subjects in the right caudate nucleus ($W = 6.99$ ($df = 2$), p -value = 0.0303), left middle temporal gyrus ($W = 6.46$ ($df = 2$), p -value = 0.0396), and right middle temporal gyrus ($W = 6.41$ ($df = 2$), p -value = 0.0406). This same analysis established differences in pattern and acceleration of decline between AD females and control females in the right caudate nucleus ($W = 6.14$ ($df = 2$), p -value = 0.0464), left thalamus ($W = 8.48$ ($df = 2$), p -value = 0.0144), and right middle temporal gyrus ($W = 5.63$ ($df = 2$), p -value = 0.0598). Additionally, this only-female analysis revealed differences in atrophy patterns between female AD individuals and female aMCI individuals in the right and left middle temporal gyrus ($W = 6.28$ ($df = 2$), p -value = 0.0432 and $W = 9.74$ ($df = 2$), p -value = 0.0077, respectively), and left thalamus ($W = 7.7$ ($df = 2$), p -value = 0.0213). A summary of these results is presented in Table 3. Mean profiles of fitted global volume measures for



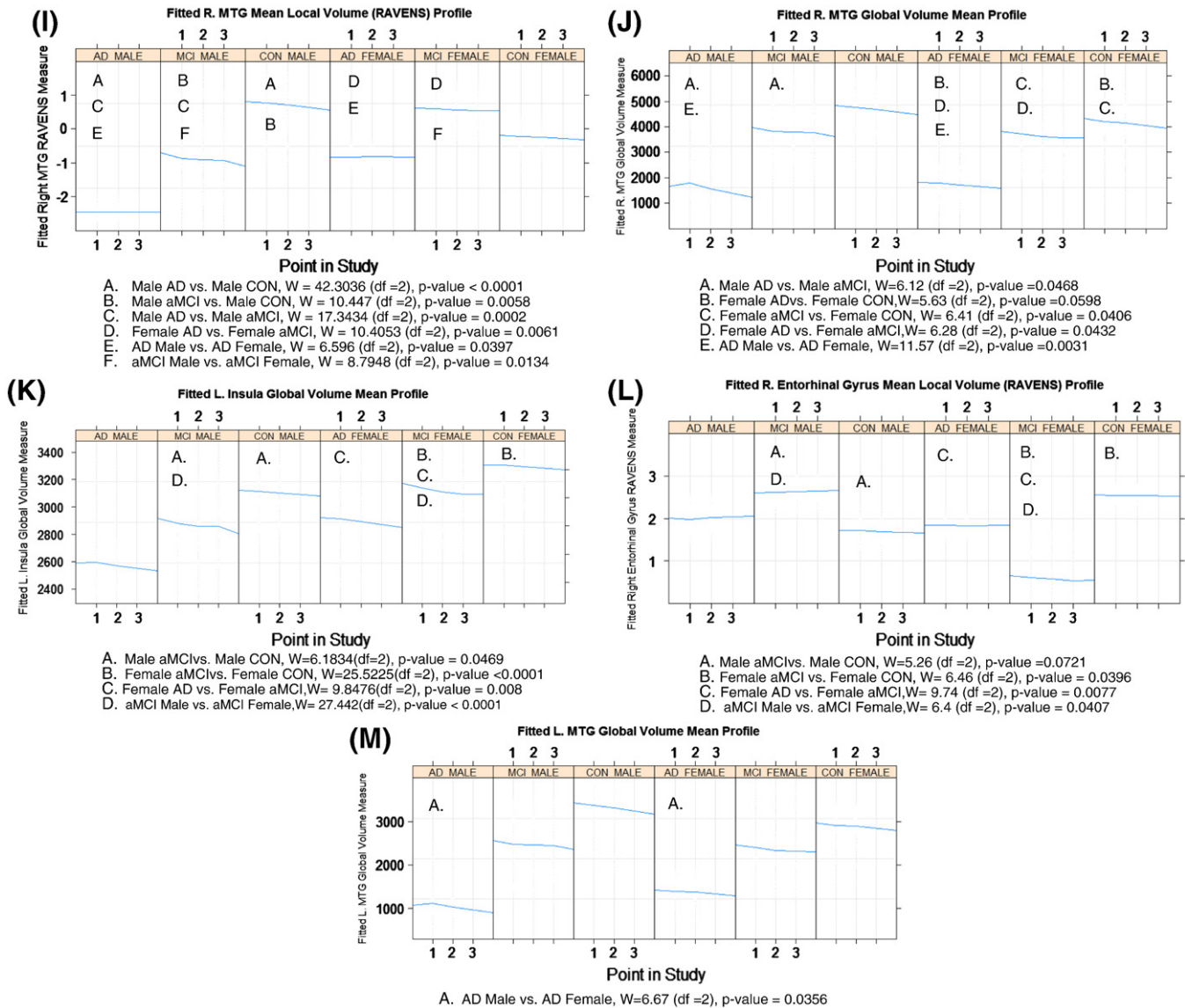


Fig. 3. Fitted mean brain volumes, corrected for ICV, for each of the relevant regions of interest over time, by sex and diagnostic group, from the global and local analyses. To correct for ICV, fitted values were corrected as follows: corrected value = fitted value $- \beta^*ICV$ (where β is the estimated coefficient from the estimated model corresponding to ICV). In the global analysis plots, the fitted value represents the estimated volume of that region based on the global analysis GEE model. In the RAVENS analysis plots, the fitted value represents the estimated RAVENS measure of the voxel based on the local analysis MAGEE model. The x-axis represents time and the y-axis represents the corrected fitted mean brain volumes as defined by the equation. Significant differences from secondary analyses are noted. Regions include (A) right precuneus (local analysis); (B) left precuneus (global analysis); (C) right caudate (global analysis); (D) left caudate (global analysis); (E) left thalamus (local analysis); (F) left thalamus (global analysis); (G) right thalamus (local analysis); (H) right amygdala (local analysis); (I) right middle temporal gyrus (local analysis); (J) right middle temporal gyrus (global analysis); (K) right entorhinal gyrus (local analysis); (L) left middle temporal gyrus (global analysis); (M) left insula (global analysis).

each region determined by the Wald statistic to be significant in the female-only analysis are shown in Fig. 3.

When the sample was restricted to include only males, results demonstrated marginally significant differences in volumetric decline patterns of the AD group compared to control males in the left precuneus ($W = 5.37$ ($df = 2$), p -value $= 0.0681$) and left thalamus ($W = 5.16$ ($df = 2$), p -value $= 0.0759$) as well as marginally significant differences between the aMCI group compared to controls in the left middle temporal gyrus ($W = 5.26$ ($df = 2$), p -value $= 0.0721$). Males with AD differed in atrophy patterns from males with aMCI in the left precuneus ($W = 10.77$ ($df = 2$), p -value $= 0.0046$) and right middle temporal gyrus ($W = 6.12$ ($df = 2$), p -value $= 0.0468$). A summary of these results is presented in Table 3. Mean profiles of fitted global volume measures for each region determined by the Wald statistic to be significant in this male-only analysis are shown in Fig. 3.

Secondary global analyses: diagnostic groups

An analysis that included only aMCI subjects revealed differences between males and females with aMCI in the pattern and acceleration of decline in the left caudate nucleus ($W = 6.79$ ($df = 2$), p -value $= 0.0336$), right caudate nucleus ($W = 8.47$ ($df = 2$), p -value $= 0.0144$), left precuneus ($W = 9.77$ ($df = 2$), p -value $= 0.0076$) and left middle temporal gyrus ($W = 6.4$ ($df = 2$), p -value $= 0.0407$). Additionally, males and females with AD were found to have differences in their pattern of brain volume decline over time in the left insula ($W = 6.67$ ($df = 2$), p -value $= 0.0356$) and right middle temporal gyrus ($W = 11.57$ ($df = 2$), p -value $= 0.0031$). Finally, an analysis of only control subjects showed differences in brain volume over time between control males and control females in the right caudate nucleus ($W = 9.3$ ($df = 2$), p -value $= 0.0096$). A summary of these results is presented in Table 4. Mean profiles of fitted

Table 4

Results from three secondary analyses, using an AD-only subset (displayed at the top), aMCI-only subset (displayed in the middle), and a control-only subset (displayed at the bottom) of the data, respectively. Only significant Wald statistics, which correspond to a test of sex effect over time, are shown for relevant regions of interest. In the "Analysis" column of the table, "local" refers to the local RAVENS map MAGEE analysis and "global" refers to the global ROI volume GEE analysis.

Male vs. Female				
Region of interest	Wald stat ($df=2$)	p-value	Analysis	Pattern
AD only				
L. insula	6.67	0.0356	Global	Male < female
L. thalamus	24.0964	<0.0001	Local	Male < female
R. middle temporal gyrus	11.57	0.0031	Global	Male < female
R. middle temporal gyrus	6.596	0.0397	Local	Male < female
R. thalamus	11.0063	0.0049	Local	Male < female
L. caudate nucleus	6.79	0.0336	Global	Male < female
L. precuneus	9.77	0.0076	Global	Female < male
L. middle temporal gyrus	6.4	0.0407	Global	Female < male
L. thalamus	25.0305	<0.0001	Local	Male < female
R. caudate nucleus	8.47	0.0144	Global	Male < female
R. entorhinal gyrus	27.442	<0.0001	Local	Female < male
R. middle temporal gyrus	8.7948	0.0134	Local	Male < female
R. precuneus	9.5174	0.0097	Local	Female < male
R. thalamus	25.7772	<0.0001	Local	Male < female
L. thalamus	9.5174	0.0097	Local	Male < female
R. amygdala	8.8576	0.0133	Local	Female < male
R. caudate nucleus	9.3	0.0096	Global	Male < female
R. precuneus	20.2069	0.0001	Local	Male < female

global volume measures for each region determined by the Wald statistic to be significant in this diagnosis-only analysis are shown in Fig. 3.

Discussion

The aim of this study was to assess whether differences exist in atrophic patterns of male and female subjects over time with probable AD and aMCI compared to elderly subjects without cognitive impairment for various grey matter structures. To do so, we utilized longitudinal ADNI data, where participants contributed an average of 3 scans to the dataset, and we analyzed these data using two analyses of local and global volume measures. We replicated prior probable AD and aMCI research findings, which previously relied on datasets that combined the sexes, by confirming those findings in more focused samples and demonstrating these results to be gender neutral. Further, our study had three novel contributions to the AD and aMCI body of research: (1) females, not males, with probable AD and aMCI differed from control counterparts in right caudate nucleus atrophy; (2) aMCI males and females differed from probable AD males and females in atrophic patterns in the thalamus, the precuneus, and the amygdala; and (3) sex differences among probable AD patients were found in the left insula, the bilateral thalamus, the right middle temporal gyrus and among the aMCI patients in the bilateral caudate nucleus, the bilateral thalamus, the bilateral precuneus, and the bilateral middle temporal gyrus.

AD and aMCI vs. controls

Our results indicate that males and females with probable AD showed significantly different patterns of decline over time than their healthy control counterparts in the right amygdala, the precuneus (bilateral for males and right for females), the bilateral thalamus, and the right middle temporal gyrus. In each of these regions, the probable AD group had markedly decreased volumes over time than the healthy controls. Similarly, compared to healthy individuals, patients with aMCI showed lower volumes over time in the right amygdala and the bilateral middle temporal gyrus. aMCI inflicted males and females also differed from their control counterparts in the right precuneus. However, while aMCI females expectedly showed patterns of less

volume over time from the controls, control males, on the other hand, showed less volume over time compared to aMCI males. Furthermore, our results revealed that females with probable AD and aMCI varied in patterns of right caudate nucleus atrophy over time compared to female healthy controls while males with AD and aMCI did not show statistically significant differences when compared to healthy males.

Our amygdala results are consistent with previous ROI and voxel-based morphometry (VBM) studies that have detected significant grey matter reduction in the bilateral amygdala in AD subjects compared to controls (Basso et al., 2006; Cuenod et al., 1993; Fennema-Notestine et al., 2009; Fjell et al., 2009; Krasuski et al., 1998; Laakso et al., 1995; Lehericy et al., 1994; Maunoury et al., 1996; Mauri et al., 1998; Mizuno et al., 2000) as well as in MCI subjects compared to controls (Fennema-Notestine et al., 2009; Krasuski et al., 1998; Lehericy et al., 1994; Mizuno et al., 2000). The amygdala has been historically implicated in the neurobiology of emotion for memory consolidation via the enhancement of memory by emotion (Buchanan and Adolphs, 2004; Labar and Cabeza, 2006; LaBar and Phelps, 1998; McGaugh et al., 1996; Phelps, 2004). Additional research supporting the amygdala's role in memory deficits has shown that amygdala lesions produce declarative memory impairment in humans (Markowitsch et al., 1994; Tranel and Hyman, 1990). Thus, our results are not surprising and confirm previous reports of reduced amygdala volumes over time in the AD and aMCI groups compared to controls, lending further support of differences of the AD and aMCI groups from healthy populations. Our study demonstrates that this finding is gender-neutral and occurs in both males and females.

Prior research has shown pronounced volume reduction of the thalamus in aMCI and in AD patients compared to individuals without cognitive deficits (Callen et al., 2001; Chen and Herskovits, 2006; Chetelat et al., 2005; de Jong et al., 2008; Good et al., 2002; Jernigan et al., 2001; Karas et al., 2004; de Oliveira et al., 2010), a result that was partially replicated. While not much is known about the exact role that the thalamus plays in cognitive processes, the pathology of the thalamus has been increasingly described in imaging studies, and its dysfunction may contribute to cognitive decline, particularly in memory performance (Aggleton and Brown, 1999). As such, the difference in the AD patients' longitudinal atrophic patterns compared to controls was not unforeseen. Research has found that the thalamus is essential for generating attention (Newman, 1995), its atrophy in AD patients is correlated with verbal learning (Desgranges et al., 1998; Stout et al., 1999), its anterior and medial nuclei are involved in declarative memory functioning (Van der Werf et al., 2000), and patients that sustain thalamic lesions also experience anterograde and remote memory loss (Stuss et al., 1988). Although we did not find differences between individuals in the aMCI group and controls in volumetric decline as other research has found, this could be because our stable aMCI group excluded patients that had converted to probable AD status during the course of the study. Teipel and colleagues hypothesized that thalamic atrophy may have some prognostic value in the conversion from aMCI to AD (Teipel et al., 2007), suggesting that differences in thalamus volume of aMCI patients compared to controls may rely on aMCI patients eventually converting to AD.

Further, we replicated previous middle temporal gyrus findings that AD patients and aMCI patients show volume reduction compared to healthy individuals in the middle temporal gyrus (Apostolova et al., 2007; Bell-McGinty et al., 2005; Chetelat et al., 2005; Desikan et al., 2008; Frisoni et al., 2002; Galton et al., 2001; Holland et al., 2009; Rabinovici et al., 2007; Schroeter et al., 2009). Previous functional neuroimaging studies have suggested that the middle temporal gyrus is involved in several cognitive processes including language and semantic memory processing (Cabeza and Nyberg, 2000; Chao et al., 1999; Tranel et al., 1997). Because both sexes in the AD and aMCI group showed these differences, this

finding suggests that the neuropathological progression of middle temporal gyrus atrophy in these two patient groups might be independent of gender.

Our caudate nucleus results are consistent with a number of volumetric MRI studies that have reported progressive atrophy of the caudate nucleus in AD and aMCI patients (Barber et al., 2002; Frisoni et al., 2002; Good et al., 2002; Madsen et al., 2010) as well as with research showing amyloid plaques and neurofibrillary tangles in the striatum of AD patients (Braak and Braak, 1990). The caudate nucleus is thought to be an important mediator in spatial working memory (Goldman and Enger Rosvold, 1972; Levy et al., 1997), which plays a significant role in planning and executing motor action in the service of behavioral objectives (Postle and D'Esposito, 1999). Few previous comparative studies have addressed the question of sex differences in caudate nucleus atrophy of cognitively impaired individuals, thus, our finding that it is females in the AD and aMCI groups, not males, that differ from their control counterparts is novel and suggests that females were more likely to express abnormal progression of caudate nucleus volume reduction than males.

Prior work suggests that AD patients have greater bilateral precuneus atrophy than controls (Driscoll et al., 2009; Frisoni et al., 2002; Good et al., 2002; Karas et al., 2007; Risacher et al., 2009) as do aMCI patients compared to controls (Risacher et al., 2009). Our findings in this region for AD males and females as well as aMCI females are in line with previous studies. However, our divergent finding that males with aMCI showed larger volumes over time than controls is somewhat puzzling. In older individuals without dementia, research indicates that longitudinal decreases in grey matter volumes, such as in the precuneus, are widespread and observed even in very healthy subjects during normal aging (Resnick et al., 2003; Thambisetty et al., 2010). This indicates that even elderly controls may show signs of distinct anatomical changes of brain regions at risk of functional decline. Further, results from a longitudinally assessment of structural changes within the aMCI group using VBM suggest that aMCI patients who eventually convert to probable AD show more overall atrophy and focal atrophy in the precuneus than non-converters (Chetelat et al., 2005). These past research studies point to a possible explanation of our unexpected result: perhaps the male controls in our sample showed enough volumetric decline in the right precuneus, as suggested by previous research, that, when compared with our sample of stable aMCI patients, it was enough to surpass them. It could be the case that by only considering those aMCI patients that did not convert, the differentiation between aMCI patients and controls, at least in the male subgroup, was opposite from expected.

AD vs. aMCI

Previous research has demonstrated that AD patients show less grey matter volume over time compared to individuals with aMCI (Chetelat et al., 2002; Fennema-Notestine et al., 2009; Madsen et al., 2010; Risacher et al., 2009), although none of these studies investigated the potential impact of sex on differences in grey matter volume reduction between these two groups. Our study revealed that males and females with probable AD showed varying atrophic patterns compared to the aMCI patients in the bilateral thalamus, the middle temporal gyrus (bilateral for females and right for males), as well as the left insula and the right amygdala (where, for both of these latter regions, differences were significant for females). Males with AD also differed from males with aMCI in the left precuneus, while females in the two groups showed a trend for a difference in this region.

Few studies have directly compared AD and aMCI patients in brain atrophic patterns in the regions that we identified as evidencing a sex-diagnosis interaction over time. Using VBM, Chetelat and colleagues found significantly greater atrophy in the right middle temporal gyrus in patients with mild AD versus aMCI (Chetelat et al., 2002) while other researchers have reported that these two groups differ in cortical thickness of the bilateral middle temporal gyrus (Risacher

et al., 2009). In the present study, we add to this finding and show that stable AD patients differ in their pattern of middle temporal gyrus atrophy from stable aMCI patients in both sex groups. In another investigation, Fennema-Notestine et al. (2009) explored pairwise comparisons of reduced amygdala volumes in AD individuals versus both those with single- and multi-domain aMCI, reporting differences among all groups in the amygdala, although they did not describe the statistical significance of this finding. Similarly, Risacher and colleagues found that AD participants had significant reductions in bilateral amygdalar volumes relative to aMCI participants (Risacher et al., 2009). Our study explicitly compared the trajectories of volume reduction in aMCI to probable AD individuals in the right amygdala, demonstrating that, in females, these patient groups differed in their right amygdala atrophy over time, with AD subjects showing less amygdala volume across time points. Further, while to our knowledge no other study has shown differences in volume reduction between AD and aMCI patients in the thalamus, one study reported texture analysis differences between the groups in the left thalamus (de Oliveira et al., 2010), lending support for our finding of thalamus volume differences over time in both males and females with AD and aMCI. Research regarding differences between the two groups in precuneus volume is inconclusive, with some reporting reduced cortical thickness in the bilateral precuneus in AD patients compared to those with aMCI (Risacher et al., 2009) and others showing both significant and insignificant peaks in grey matter density reduction in the left precuneus and no difference between the groups in the right precuneus (Chetelat et al., 2002). Our findings were similarly inconclusive, demonstrating only that males with aMCI and AD differed in their patterns of left precuneus volume reduction over time. In the caudate nucleus, some researchers have reported no differences in volume between these two groups (Chetelat et al., 2002; Madsen et al., 2010), supporting our finding of a lack of difference between the groups in this region. Thus, our work demonstrates not only that these two groups differ in atrophic patterns in the thalamus, amygdala, middle temporal gyrus, insula and precuneus, but that these differences are gender-neutral in the thalamus and middle temporal gyrus, female-specific in the insula and the amygdala, and male-specific in the left precuneus.

Within group sex differences

Finally, in an exploration of within-group sex differences, we found that males showed distinctly lower volumes over time compared to females in the following regions: the left insula, the bilateral thalamus, and the right middle temporal gyrus for the AD group; the bilateral thalamus, the bilateral caudate nucleus, and the right middle temporal gyrus for the aMCI group, as well as the left thalamus, the right caudate nucleus, and the right precuneus in the control group. Conversely, the opposite pattern (where females showed decreased volumes of structures over time compared to males) was found in the left middle temporal gyrus and the bilateral precuneus for aMCI patients, and the right amygdala for controls.

Our caudate nucleus findings are supported by previous research. In healthy women compared to healthy men, imaging studies of this region have reported relatively larger caudate volumes (Filipek et al., 1994; Goldstein et al., 2001; Murphy et al., 1996) and higher caudate presynaptic dopamine synthesis capacity (Laakso et al., 2002) in women. In an analysis that included healthy control, AD patients, and subjects with aMCI, Madsen et al. (2010) reported smaller caudate volumes in men than women when all subjects were combined (although this analysis did not test how sex impacted volume for each diagnostic group separately). Further, because the caudate nucleus is a grey matter structure that lies just below the boundary of the lateral ventricles, a decrease in caudate nucleus volume typically coincides with an increase in lateral ventricle volume. As such, Grant et al. (1987) reported that men, but not women, exhibited a significant age-

related increase in lateral ventricular volume. Likewise, Blatter et al. (1995) observed higher correlations in men than women between age and lateral ventricle volume. The result of these studies coincide with our finding of sex differences among control subjects in caudate nucleus volume and our investigation introduces the finding of sex differences among the aMCI group as a novel contribution.

There is limited research concerning sex differences in thalamic volume reduction. In healthy controls, while some research has found that thalamic volume declined linearly with age at a similar rate in both men and women (Sullivan et al., 2004), others have reported that brain matter area in the thalamus was significantly larger in female subjects than male subjects (Xu et al., 2000). In an investigation of a patient population of Japanese schizophrenics (Yotsutsuji et al., 2003), volumes of the third ventricle, a cavity in the brain which the thalamus surrounds, were larger in the patients compared to the control subjects, and larger in the males than in the females. This research suggests support for our findings, which indicated that males have markedly less volume over time than females in the bilateral thalamus for the AD and aMCI groups and the right thalamus for the control groups. Further, sex differences in patterns of atrophy of this region were not necessarily specific to diagnosis.

The difference between control males and control females in precuneus volume reduction has been seen in previous investigations. In a study of anatomical changes in regional grey matter volumes in a group of older adults without dementia, males showed greater rates of decline in the right precuneus than females (Thambisetty et al., 2010), just as we found in our secondary analyses which included only control subjects. Our results also revealed differences between the sexes in the aMCI group. However, these results showed the inverse to what we saw in the healthy control analysis such that females exhibited less volume over time than males. No previous research has investigated the specific question of how sex impacts precuneus atrophy among AD or aMCI participants. Our finding that females in the aMCI group compared to males showed less precuneus volume over time suggests the possibility of gender-specific abnormalities in the precuneus and that the pathology in aMCI has greater predilection for the precuneus in males than females. More research is needed in this area.

Research focusing on sex differences in insular, amygdala, and middle temporal gyrus atrophy is limited and inconclusive. Our result that males and females in the control group showed different patterns of amygdala decrease over time is supported by research indicating that males exhibit greater volumes or neuronal densities in the amygdala (Giedd et al., 1996; Goldstein et al., 2001; Witte et al., 2010). Further, Cahill and colleagues found functional differences between the sexes in this region, suggesting that males showed increased activity in the right amygdala related to enhanced memory for emotional stimuli (Cahill et al., 2001). However, investigations of AD and aMCI patients have yet to show men and women to have varying atrophy patterns in this region (Krasuski et al., 1998; Mizuno et al., 2000). In our study, we also did not find such sex differences in the amygdala for the aMCI and AD groups, indicating that such differences may not exist in clinical populations with cognitive impairments. Further, although other researchers have not previously investigated sex differences in insular volumes within the AD patient group, several fMRI memory studies have indicated that males and females differentially activate the insula while performing a memory related task, suggesting that there are sex-related differences in relative BOLD signal change associated with memory (Piefke et al., 2005; Valera et al., 2010). Our finding and lack of research on this topic point to a need for additional research regarding sexual dimorphism of insular atrophy in AD patients. Additionally, sex differences of volumetric reduction of the right middle temporal gyrus were observed for both the AD and aMCI groups, such that males showed reduced volume across time compared to AD and aMCI females. The result for the left middle temporal gyrus in the aMCI group was the opposite, where females showed lower volumes over time. This finding is novel

because relatively few studies have investigated differences within AD and aMCI patients groups for sex differences in this region.

Declining trend in volume

As depicted in Fig. 3, most of the regions identified by our primary analyses show a declining trend in mean volumetric measures over time for AD patients, subjects with aMCI, and healthy controls. There are a few unexpected exceptions that warrant comment. In particular, the mean trends shown for the middle temporal gyrus, entorhinal cortex, and precuneus do not decrease substantially in subjects with AD and aMCI, but do show decrease in control subjects. Previous research supports the findings that age has an effect on the volume of these brain structures in healthy individuals. For instance, investigations have found age-related decline in the precuneus in healthy volunteers, but not AD subjects using VBM (Ohnishi et al., 2001). In the entorhinal cortex, age was found to significantly affect the rates of atrophy in cognitively healthy subjects (Du et al., 2006). Brain volume atrophy in the middle temporal gyrus has also been reported for healthy aging subjects (Fjell et al., 2009).

Limitations

Several caveats should be considered in this work. Twenty-four ROIs were selected for investigation in the present analyses. It is possible that choice of alternative or additional ROIs could reveal other differences between groups. A further limitation is the uneven sex balance across groups in the primary sample, with males being somewhat over-represented among aMCI patients. The full ADNI cohort shows a similar bias toward greater representation of males, particularly in the aMCI group.

Additionally, our analysis utilized data from stable AD subjects who were relatively early in their cognitive decline as well as from individuals with stable aMCI who were not separated into single- and multi-domain subtypes. Thus, because the ADNI dataset does not represent the population at large, these results would not apply to persons in very advanced stages of Alzheimer's. Our intention behind separating all the groups into converters and non-converters follows the example of previous ADNI publications (e.g., Risacher et al., 2009). In the present study, we considered only stable aMCI subjects due to our confidence in discriminating between stable AD and stable aMCI patients and between stable aMCI patients and healthy controls in the pre-selected ROIs. Our research question of interest was motivated by whether and how sex plays a role in atrophy between these groups, not necessarily how sex fits into conversion between diagnoses. Further, the aMCI converter group was homogenous in how much study time it took for subjects to convert (while enrolled in the study). In our investigation, we determined that such variability might complicate our model to the extent that interpreting findings would be difficult without first having a sense of how the stable groups compared. We acknowledge that including aMCI converters as a separate group and further classifying the aMCI patients in domain categories may be a more clinically relevant question than the discrimination between aMCI non-converters and healthy controls, and should be pursued in future research.

Further, we examined atrophy in grey matter volume in our sample using both the traditional ROI-based methods and the newer RAVENS voxel-wise methods. While we saw similar trends between two methods in patients diagnosed with AD, subjects with aMCI, and healthy controls in the middle temporal gyrus and the thalamus, there was a lack of correspondence of significant findings between the two methods in the other regions. The primary cause for the discrepancy between the voxel-wise RAVENS and the ROI results likely stems from the methodological differences between voxel-averaged, landmark-based ROI analyses and the single, voxel-by-voxel whole brain RAVENS measurements. The ROI method is based on the *a priori* definition of a number of regions of interest, followed by measurements of tissue

volumes within each ROI. The RAVENS method, on the other hand, performs spatial normalization of volumetric images into a stereotaxic space, followed by a subsequent voxel-wise statistical analysis of the resulting spatial distributions of grey matter. As evidenced by our findings, it seems that these two approaches do not measure the exact same entities and it is not yet clear whether one is more reliable for reflecting volume trajectories than the other. However, because both methods provide different types of information, they should be used in tandem.

Some potential confounds not included in the current analysis should also be mentioned and considered in future studies. Recent research has shown that a history of depression, anxiety, or substance abuse may adversely affect the grey matter volumes of certain brain structures of AD and aMCI patients, in particular, the hippocampus, entorhinal cortex, or amygdala (Bell-McGinty et al., 2002; Rapp et al., 2006; Sheline et al., 2003). While subjects in our sample did not have major depression, a history of alcohol or substance abuse, nor a dependency on alcohol or another substance within the past 2 years due to exclusion criteria for ADNI, we did not have data regarding history of clinically relevant depression or anxiety. It is possible that if this information was available and included in our model, this could have altered our results for the hippocampus, entorhinal cortex, or amygdala. Our findings, therefore, need to be interpreted with caution and with the knowledge that history of depression or anxiety was not considered during modeling. Future studies, which account for these variables, will need to be undertaken.

Contributions

In this analysis, we utilized two measures to assess volumetric decline of certain grey matter structures: the RAVENS measure allowed for the consideration of local changes while the global ROI measure permitted investigation of spatially defined, whole brain regions. The conclusions that we drew from the results of the local and global analyses in the left thalamus and the right middle temporal gyrus overlapped, but both methods additionally contributed unique information to our conclusions, suggesting that sole measurements are not sufficient for comparing AD, aMCI, and controls to each other. Further, both the MAGEE and GEE analysis techniques provided more advanced longitudinal statistical designs than those used in previous longitudinal imaging studies of these subject groups (e.g., Hua et al., 2010) and allowed for the inclusion of more than two scans in the analyses. These methods complement each other and, when coupled, can be used as parallel tools to provide different insights into data.

Our findings also emphasize the importance of considering sex when characterizing changes in brain structure in AD and aMCI patients. If grey matter volume change differs between female and male patients in certain brain regions, then this should be taken into account when determining disease progression and predicting likelihood of conversion of aMCI to AD. Further, in the future, sex group comparisons between patients and controls may be needed to improve the sensitivity and specificity of grey matter volumetry in the evaluation of patients with AD and aMCI.

Conclusion

In this longitudinal investigation of the changes in brain structure volumes, we conclude that sex differences among AD and aMCI participants are an important consideration in determining how groups differ from each other in structural atrophy. In the right caudate nucleus, females with AD and aMCI showed differences from controls that males did not. Further, AD males and females showed different patterns of volumetric decline compared to aMCI counterparts in the thalamus, precuneus, and amygdala. Additionally, sex differences among individuals in the patient groups were found in the insula, thalamus, middle temporal

gyrus, caudate nucleus, and precuneus. To our knowledge, these are the first reported results that use more than two time points to assess longitudinal volumetric decline differences among sex and diagnostic groups. Future studies are needed to replicate and extend current findings.

Acknowledgments

M. Skup is supported in part by training grant T32 MH014235 from the National Institute on Mental Health. H. Zhu is supported in part by NSF grant BCS-08-26844 and NIH grants RR025747-01, P01CA142538-01, MH086633, and AG03338. Y. Fan is supported in part by the National Science Foundation of China grant 30970770 and by the Hundred Talents Programs, Chinese Academy of Sciences. W. Lin is supported in part by NIH grants R01NS055754 and R01EB5-34816.

Data collection and sharing for this project were funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, Astra Zeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc., F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

References

- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), *Proceedings of the Second International Symposium on Information Theory*. Akademiai Kiado, Budapest, pp. 267–281.
- Apostolova, L., Dinov, I., Dutton, R., Hayashi, K., Toga, A., Cummings, J., Thompson, P., 2006. 3D comparison of hippocampal atrophy in amnesic mild cognitive impairment and Alzheimer's disease. *Brain* 129 (11), 2867.
- Apostolova, L., Steiner, C., Akopyan, G., Dutton, R., Hayashi, K., Toga, A., Cummings, J., Thompson, P., 2007. Three-dimensional gray matter atrophy mapping in mild cognitive impairment and mild Alzheimer disease. *Arch. Neurol.* 64 (10), 1489.
- Aggleton, J.P., Brown, M.W., 1999. Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behav. Brain Sci.* 22 (3), 425.
- Allen, J., Bruss, J., Brown, C., Damasio, H., 2005. Normal neuroanatomical variation due to age: the major lobes and a parcellation of the temporal region. *Neurobiol. Aging* 26 (9), 1245–1260.
- Bai, F., Zhang, Z., Watson, D., Yu, H., Shi, Y., Zhu, W., Wang, L., Yuan, Y., Qian, Y., 2009. Absent gender differences of hippocampal atrophy in amnesic type mild cognitive impairment. *Neurosci. Lett.* 450 (2), 85–89.
- Ballmaier, M., O'Brien, J., Burton, E., Thompson, P., Rex, D., Narr, K., McKeith, I., DeLuca, H., Toga, A., 2004. Comparing gray matter loss profiles between dementia with Lewy bodies and Alzheimer's disease using cortical pattern matching: diagnosis and gender effects. *Neuroimage* 23 (1), 325–335.
- Barber, R., McKeith, I., Ballard, C., O'Brien, J., 2002. Volumetric MRI study of the caudate nucleus in patients with dementia with Lewy bodies, Alzheimer's disease, and vascular dementia. *J. Neurol. Neurosurg. Psychiatry* 72 (3), 406.
- Barnes, L., Wilson, R., Schneider, J., Bienias, J., Evans, D., Bennett, D., 2003. Gender, cognitive decline, and risk of AD in older persons. *Neurology* 60, 1777–1781.
- Basso, M., Yang, J., Warren, L., MacAvoy, M., Varma, P., Bronen, R., van Dyck, C., 2006. Volumetry of amygdala and hippocampus and memory performance in Alzheimer's disease. *Psychiatry Res. Neuroimaging* 146 (3), 251–261.
- Bayles, K., Azuma, T., Cruz, R., Tomoeda, C., Wood, J., Montgomery Jr., E., 1999. Gender differences in language of Alzheimer disease patients revisited. *Alzheimer Dis. Assoc. Disord.* 13 (3), 138.
- Bell-McGinty, S., Butters, M., Meltzer, C., Greer, P., Reynolds, C., Becker, J., 2002. Brain morphometric abnormalities in geriatric depression: long-term neurobiological effects of illness duration. *Am. J. Psychiatry* 159, 1424–1427.

- Bell-McGinty, S., Lopez, O., Meltzer, C., Scanlon, J., Whyte, E., DeKosky, S., Becker, J., 2005. Differential cortical atrophy in subgroups of mild cognitive impairment. *Arch. Neurol.* 62 (9), 1393.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Methodol.* 57 (1), 289–300.
- Beresford, T., Arciniegas, D., Alfors, J., Clapp, L., Martin, B., Du, Y., Liu, D., Shen, D., Davatzikos, C., 2006. Hippocampus volume loss due to chronic heavy drinking. *Alcohol. Clin. Exp. Res.* 30 (11), 1866–1870.
- Blatter, D.D., Bigler, E.D., Gale, S.D., Johnson, S.C., Anderson, C.V., Burnett, B.M., Parker, N., Kurth, S., Horn, S.D., 1995. Quantitative volumetric analysis of brain MR: normative database spanning 5 decades of life. *Am. J. Neuroradiol.* 16, 241–251.
- Braak, H., Braak, E., 1990. Alzheimer's disease: striatal amyloid deposits and neurofibrillary changes. *J. Neuropathol. Exp. Neurol.* 49 (3), 215.
- Brun, C., Lepore, N., Luders, E., Chou, Y., Madsen, S., Toga, A., Thompson, P., 2009. Sex differences in brain structure in auditory and cingulate regions. *NeuroReport* 20 (10), 930.
- Buchanan, T., Adolphs, R., 2004. The Neuroanatomy of Emotional Memory in Humans.
- Buckwalter, J., Sobel, E., Dunn, M., Diz, M., Henderson, V., 1993. Gender differences on a brief measure of cognitive functioning in Alzheimer's disease. *Arch. Neurol.* 50 (7), 757.
- Cabeza, R., Nyberg, L., 2000. Imaging cognition II: an empirical review of 275 PET and fMRI studies. *J. Cogn. Neurosci.* 12 (1), 1–47.
- Cahill, L., Haier, R., White, N., Fallon, J., Kilpatrick, L., Lawrence, C., Potkin, S., Alkire, M., 2001. Sex-related difference in amygdala activity during emotionally influenced memory storage. *Neurobiol. Learn. Mem.* 75 (1), 1–9.
- Callen, D., Black, S., Caldwell, C., Grady, C., 2004. The influence of sex on limbic volume and perfusion in AD. *Neurobiol. Aging* 25 (6), 761–770.
- Callen, D., Black, S., Gao, F., Caldwell, C., Szalai, J., 2001. Beyond the hippocampus: MRI volumetry confirms widespread limbic atrophy in AD. *Neurology* 57 (9), 1669.
- Carmichael, O., Kuller, L., Lopez, O., Thompson, P., Dutton, R., Lu, A., Lee, S., Lee, J., Aizenstein, H., Meltzer, C., Liu, Y., Toga, A.W., Becker, J.T., 2007. Ventricular volume and dementia progression in the Cardiovascular Health Study. *Neurobiol. Aging* 28 (3), 389–397.
- Chao, L., Haxby, J., Martin, A., 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat. Neurosci.* 2, 913–919.
- Chen, R., Herskovits, E., 2006. Network analysis of mild cognitive impairment. *NeuroImage* 29 (4), 1252–1259.
- Chetelat, G., Desgranges, B., de la Sayette, V., Viader, F., Eustache, F., Baron, J., 2002. Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment. *NeuroReport* 13, 1939–1943.
- Chetelat, G., Landeau, B., Eustache, F., Mezenge, F., Viader, F., de la Sayette, V., Desgranges, B., Baron, J., 2005. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: a longitudinal MRI study. *NeuroImage* 27 (4), 934–946.
- Cockrell, J., Folstein, M., 1988. Mini-Mental State Examination (MMSE). *Psychopharmacol. Bull.* 24 (4), 689.
- Coffey, C., Wilkinson, W., Parashos, L., Soady, S., Sullivan, R., Patterson, L., Figiel, G., Webb, M., Spritzer, C., Djang, W., 1992. Quantitative cerebral anatomy of the aging human brain: a cross-sectional study using magnetic resonance imaging. *Neurology* 42 (3), 527.
- Cowell, P., Turetsky, B., Gur, R., Grossman, R., Shtasel, D., Gur, R., 1994. Sex differences in aging of the human frontal and temporal lobes. *J. Neurosci.* 14, 4748–4755.
- Cuenod, C., Denys, A., Michot, J., Jehenson, P., Forette, F., Kaplan, D., Syrota, A., Boller, F., 1993. Amygdala atrophy in Alzheimer's disease: an in vivo magnetic resonance imaging study. *Arch. Neurol.* 50 (9), 941.
- Davatzikos, C., 1998. Mapping image data to stereotaxic spaces: applications to brain mapping. *Hum. Brain Mapp.* 6 (5–6), 334–338.
- Davatzikos, C., Genc, A., Xu, D., Resnick, S., 2001. Voxel-based morphometry using the RAVENTS maps: methods and validation using simulated longitudinal atrophy. *NeuroImage* 14 (6), 1361–1369.
- De Jong, L., van der Hiele, K., Veer, I., Houwing, J., Westendorp, R., Bollen, E., de Bruin, P., Middelkoop, H., van Buchem, M., van der Grond, J., 2008. Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. *Brain* 131 (12), 3277.
- de Oliveira, M., Balthazar, M., D'Abreu, A., Yasuda, C., Damasceno, B., Cendes, F., Castellano, G., 2010. MR imaging texture analysis of the corpus callosum and thalamus in amnesic mild cognitive impairment and mild Alzheimer disease. *Am. J. Neuroradiol.* 32 (1), 60–66.
- Desgranges, B., Baron, J., de la Sayette, V., Petit-Taboue, M., Benali, K., Landeau, B., Lechevalier, B., Eustache, F., 1998. The neural substrates of memory systems impairment in Alzheimer's disease. A PET study of resting brain glucose utilization. *Brain* 121 (4), 611.
- Desikan, R., Fischl, B., Cabral, H., Kemper, T., Guttman, C., Blacker, D., Hyman, B., Albert, M., Killiany, R., 2008. MRI measures of temporoparietal regions show differential rates of atrophy during prodromal AD. *Neurology* 71 (11), 819.
- Drachman, D., Swearer, J., O'Donnell, B., Mitchell, A., Maloon, A., 1992. The Caretaker Obstreperous-Behavior Rating Assessment (COBRA) scale. *J. Am. Geriatr. Soc.* 40 (5), 463.
- 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.
- Du, A., Schuff, N., Chao, L., Kornak, J., Jagust, W., Kramer, J., Reed, B., Miller, B., Norman, D., Chui, H., Weiner, M., 2006. Age effects on atrophy rates of entorhinal cortex and hippocampus. *Neurobiol. Aging* 27 (5), 733–740.
- Dubois, B., Feldman, H., Jacova, C., DeKosky, S., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P.J., Scheltens, P., 2007. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol.* 6 (8), 734–746.
- Edland, S., Xu, Y., Plevak, M., O'Brien, P., Tangalos, E., Petersen, R., Jack Jr., C., 2002. Total intracranial volume: normative values and lack of association with Alzheimer's disease. *Neurology* 59 (2), 272.
- Fan, Y., Batmanghelich, N., Clark, C., Davatzikos, C., 2008. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *NeuroImage* 39 (4), 1731–1743.
- Fennema-Notestine, C., Hagler Jr., D., McEvoy, L., Fleisher, A., Wu, E., Karow, D., Dale, A., 2009. Structural MRI biomarkers for preclinical and mild Alzheimer's disease. *Hum. Brain Mapp.* 30 (10), 3238–3253.
- Filipek, P., Richelme, C., Kennedy, D., Caviness Jr., V., 1994. The young adult human brain: an MRI-based morphometric analysis. *Cereb. Cortex* 4 (4), 344.
- Fjell, A., Walhovd, K., Fennema-Notestine, C., McEvoy, L., Hagler, D., Holland, D., Brewer, J., Dale, A., 2009. One-year brain atrophy evident in healthy aging. *J. Neurosci.* 29 (48), 15223.
- Fleisher, A., Grundman, M., Jack Jr., C., Petersen, R., Taylor, C., Kim, H., Schiller, D., Bagwell, V., Sencakova, D., Weiner, M., DeCarli, C., DeKosky, S.T., van Dyck, C.H., Thal, L.J., Study, Alzheimer's Disease Cooperative, 2005. Sex, apolipoprotein E ϵ 4 status, and hippocampal volume in mild cognitive impairment. *Arch. Neurol.* 62 (6), 953.
- Folstein, M., Folstein, S., McHugh, P., Fanjiang, G., 1975. Mini-Mental State Examination (MMSE). *J. Psychiatr. Res.* 12, 189–198.
- Freeborough, P., Fox, N., Kitney, R., 1997. Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. *Comput. Meth. Programs Biomed.* 53 (1), 15–25.
- Frisoni, G., Testa, C., Zorzan, A., Sabattoli, F., Beltramello, A., Soininen, H., Laakso, M., 2002. Detection of grey matter loss in mild Alzheimer's disease with voxel-based morphometry. *J. Neurol. Neurosurg. Psychiatry* 73 (6), 657.
- Friston, K., Holmes, A., Poline, J., Grasby, P., Williams, S., Frackowiak, R., Turner, R., 1995. Analysis of fMRI time-series revisited. *NeuroImage* 2 (1), 45–53.
- Galton, C., Patterson, K., Graham, K., Lambon-Ralph, M., Williams, G., Antoun, N., Sahakian, B., Hodges, J., 2001. Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology* 57 (2), 216.
- Gao, S., Hendrie, H., Hall, K., Hui, S., 1998. The relationships between age, sex, and the incidence of dementia and Alzheimer disease: a meta-analysis. *Arch. Gen. Psychiatry* 55 (9), 809.
- Giedd, J., Snell, J., Lange, N., Rajapakse, J., Casey, B., Kozuch, P., Vaituzis, A., Vauss, Y., Hamburger, S., Kaysen, D., Rapoport, J.L., 1996. Quantitative magnetic resonance imaging of human brain development: ages 4–18. *Cereb. Cortex* 6 (4), 551.
- Goldman, P., Enger Rosvold, H., 1972. The effects of selective caudate lesions in infant and juvenile rhesus monkeys. *Brain Res.* 43 (1), 53–66.
- Goldstein, J., Seidman, L., Horton, N., Makris, N., Kennedy, D., Caviness, V., Faraone, S., Tsuang, M., 2001. Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb. Cortex* 11 (6), 490.
- Goldszal, A., Davatzikos, C., Pham, D., Yan, M., Bryan, R., Resnick, S., 1998. An image-processing system for qualitative and quantitative volumetric analysis of brain images. *J. Comput. Assist. Tomogr.* 22 (5), 827.
- Good, C., Scallan, R., Fox, N., Ashburner, J., Friston, K., Chan, D., Crum, W., Rossor, M., Frackowiak, R., 2002. Automatic differentiation of anatomical patterns in the human brain: validation with studies of degenerative dementias. *NeuroImage* 17 (1), 29–46.
- Grant, R., Condon, B., Lawrence, A., Hadley, D., Patterson, J., Bone, L., Teasdale, G., 1987. Human cranial CSF volumes measured by MRI: sex and age influences. *Magn. Reson. Imaging* 5 (6), 465–468.
- Hebert, L., Wilson, R., Gilley, D., Beckett, L., Scherr, P., Bennett, D., Evans, D., 2000. Decline of language among women and men with Alzheimer's disease. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 55 (6), P354.
- Henderson, V., Buckwalter, J., 1994. Cognitive deficits of men and women with Alzheimer's disease. *Neurology* 44 (1), 90.
- Herholz, K., Schopphoff, H., Schmidt, M., Mielke, R., Eschner, W., Scheidhauer, K., Schicha, H., Heiss, W., Ebmeier, K., 2002. Direct comparison of spatially normalized PET and SPECT scans in Alzheimer's disease. *J. Nucl. Med.* 43 (1), 21.
- Herlitz, A., Nilsson, L., Backman, L., 1997. Gender differences in episodic memory. *Mem. Cogn.* 25, 801–811.
- Ho, A., Raji, C., Becker, J., Lopez, O., Kuller, L., Hua, X., Lee, S., Hibar, D., Dinov, I., Stein, J., Jack, C.R., Weiner, M.W., Toga, A.W., Thompson, P.M., Cardiovascular Health Study, ADNI, 2010. Obesity is linked with lower brain volume in 700 AD and MCI patients. *Neurobiol. Aging* 31 (8), 1326–1339.
- Holland, D., Brewer, J., Hagler, D., Fennema-Notestine, C., Dale, A., 2009. Subregional neuroanatomical change as a biomarker for Alzheimer's disease. *Proc. Natl Acad. Sci.* 106 (49), 20954.
- Hua, X., Hibar, D., Lee, S., Toga, A., Jack, C., Weiner, M., Thompson, P., 2010. Sex and age differences in atrophic rates: an ADNI study with $n = 1368$ MRI scans. *Neurobiol. Aging* 31 (8), 1463–1480.
- Hughes, C., Berg, L., Danziger, W., Coben, L., Martin, R., 1982. A new clinical scale for the staging of dementia. *Br. J. Psychiatry* 140 (6), 566.
- Jack Jr., C., Bernstein, M., Fox, N., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P., et al., 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27 (4), 685–691.
- Jack Jr., C., Petersen, R., Xu, Y., Waring, S., O'Brien, P., Tangalos, E., Smith, G., Ivnik, R., Kokmen, E., 1997. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology* 49 (3), 786.
- Jack Jr., C., Shiung, M., Gunter, J., O'Brien, P., Weigand, S., Knopman, D., Boeve, B., Ivnik, R., Smith, G., Cha, R., Tangalos, E.G., Petersen, R.C., 2004. Comparison of different

- MRI brain atrophy rate measures with clinical disease progression in AD. *Neurology* 62 (4), 591.
- Jack Jr., C., Shiung, M., Weigand, S., O'Brien, P., Gunter, J., Boeve, B., Knopman, D., Smith, G., Ivnik, R., Tangalos, E., Petersen, R.C., 2005. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnesic MCI. *Neurology* 65 (8), 1227.
- Jernigan, T., Archibald, S., Fennema-Notestine, C., Gamst, A., Stout, J., Bonner, J., Hesselink, J., 2001. Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiol. Aging* 22 (4), 581–594.
- Kabani, N., MacDonald, D., Holmes, C., Evans, A., 1998. A 3D atlas of the human brain. *Neuroimage* 7, 5717.
- Karas, G., Scheltens, P., Rombouts, S., van Schijndel, R., Klein, M., Jones, B., van der Flier, W., Vrenken, H., Barkhof, F., 2007. Precuneus atrophy in early-onset Alzheimer's disease: a morphometric structural MRI study. *Neuroradiology* 49 (12), 967–976.
- Karas, G., Scheltens, P., Rombouts, S., Visser, P., Van Schijndel, R., Fox, N., Barkhof, F., 2004. Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. *Neuroimage* 23 (2), 708–716.
- Kidron, D., Black, S., Stanchev, P., Buck, B., Szalai, J., Parker, J., Szekely, C., Bronskill, M., 1997. Quantitative MR volumetry in Alzheimer's disease: topographic markers and the effects of sex and education. *Neurology* 49 (6), 1504.
- Krasuski, J., Alexander, G., Horwitz, B., Daly, E., Murphy, D., Rapoport, S., Schapiro, M., 1998. Volumes of medial temporal lobe structures in patients with Alzheimer's disease and mild cognitive impairment (and in healthy controls). *Biol. Psychiatry* 43 (1), 60–68.
- Kukull, W., Higdon, R., Bowen, J., McCormick, W., Teri, L., Schellenberg, G., van Belle, G., Jolley, L., Larson, E., 2002. Dementia and Alzheimer's disease incidence: a prospective cohort study. *Arch. Neurol.* 59 (11), 1737.
- Laakso, A., Viikman, H., Bergman, J., Haaparanta, M., Solin, O., Syvalahti, E., Salokangas, R., Hietala, J., 2002. Sex differences in striatal presynaptic dopamine synthesis capacity in healthy subjects. *Biol. Psychiatry* 52 (7), 759–763.
- Laakso, M., Partanen, K., Lehtovirta, M., Hallikainen, M., Hanninen, T., Vainio, P., Riekkinen, P., Soininen, H., 1995. MRI of amygdala fails to diagnose early Alzheimer's disease. *NeuroReport* 6 (17), 2414.
- LaBar, K., Cabeza, R., 2006. Cognitive neuroscience of emotional memory. *Nat. Rev. Neurosci.* 7 (1), 54–64.
- LaBar, K., Phelps, E., 1998. Arousal-mediated memory consolidation: role of the medial temporal lobe in humans. *Psychol. Sci.* 9 (6), 490.
- Larrabee, G., Crook, T., 1993. Do men show more rapid age associated decline in simulated everyday verbal memory than do women? *Psychol. Aging* 8, 68–71.
- Lehericy, S., Baulac, M., Chiras, J., Pierot, L., Martin, N., Pillon, B., Deweer, B., Dubois, B., Marsault, C., 1994. Amygdalohippocampal MR volume measurements in the early stages of Alzheimer disease. *Am. J. Neuroradiol.* 15 (5), 929.
- Levy, R., Friedman, H., Davachi, L., Goldman-Rakic, P., 1997. Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *J. Neurosci.* 17 (10), 3870.
- Liang, K.Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73 (1), 13.
- Li, Y., Zhu, H., Lin, W., Shen, D., Ibrahim, J., 2011. MARM: Multiscale adaptive regression models for neuroimaging data. *Journal of Royal Statistical Society in Series B. In press.*
- Luders, E., Gaser, C., Narr, K., Toga, A., 2009. Why sex matters: brain size independent differences in gray matter distributions between men and women. *J. Neurosci.* 29 (45), 14265.
- Lyketsos, C., Steele, C., Galik, E., Rosenblatt, A., Steinberg, M., Warren, A., Sheppard, J., 1999. Physical aggression in dementia patients and its relationship to depression. *Am. J. Psychiatry* 156 (1), 66.
- Madsen, S., Ho, A., Hua, X., Saharan, P., Toga, A., Jack, C., Weiner, M., Thompson, P., 2010. 3D maps localize caudate nucleus atrophy in 400 Alzheimer's disease, mild cognitive impairment, and healthy elderly subjects. *Neurobiol. Aging* 31 (8), 1312–1325.
- Markowitsch, H., Calabrese, P., Wurker, M., Durwen, H., Kessler, J., Babinsky, R., Brechtelsbauer, D., Heuser, L., Gehlen, W., 1994. The amygdala's contribution to memory—a study on two patients with Urbach-Wiethe disease. *NeuroReport* 5 (11), 1349.
- Maunoury, C., Michot, J., Caillet, H., Parlato, V., Leroy-Willig, A., Jehenson, P., Syrota, A., Boller, F., 1996. Specificity of temporal amygdala atrophy in Alzheimer's disease: quantitative assessment with magnetic resonance imaging. *Dement. Geriatr. Cogn. Disord.* 7 (1), 10–14.
- Mauri, M., Sibilla, L., Bono, G., Carlesimo, G., Sinforiani, E., Martelli, A., 1998. The role of morpho-volumetric and memory correlations in the diagnosis of early Alzheimer dementia. *J. Neurol.* 245 (8), 525–530.
- Maylor, E., Reimers, S., Choi, J., Collaer, M., Peters, M., Silverman, I., 2007. Gender and sexual orientation differences in cognition across adulthood: age is kinder to women than to men regardless of sexual orientation. *Arch. Sex. Behav.* 36, 235–249.
- McAuliffe, M., Lalonde, F., McGarry, D., Gandler, W., Csaky, K., Trus, B., 2002. Medical image processing, analysis and visualization in clinical research. *Computer-Based Medical Systems*, 2001. CBMS 2001, Proceedings. 14th IEEE Symposium on. IEEE, pp. 381–386.
- McGaugh, J., Cahill, L., Roozendaal, B., 1996. Involvement of the amygdala in memory storage: interaction with other brain systems. *Proc. Natl Acad. Sci. USA* 93 (24), 13508.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology* 34, 939–944.
- Meyer, J., Rauch, G., Crawford, K., Rauch, R., Konno, S., Akiyama, H., Terayama, Y., Haque, A., 1999. Risk factors accelerating cerebral degenerative changes, cognitive decline and dementia. *Int. J. Geriatr. Psychiatry* 14, 1050–1061.
- Mizuno, K., Wakai, M., Takeda, A., Sobue, G., 2000. Medial temporal atrophy and memory impairment in early stage of Alzheimer's disease: an MRI volumetric and memory assessment study. *J. Neurol. Sci.* 173 (1), 18–24.
- Morris, J., 1993. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 43, 2412–2414.
- Mueller, E., Moore, M., Kerr, D., Sexton, G., Camicioli, R., Howieson, D., Quinn, J., Kaye, J., 1998. Brain volume preserved in healthy elderly through the eleventh decade. *Neurology* 51 (6), 1998.
- Mungas, D., 1991. In-office mental status testing: a practical guide. *Geriatrics* 46 (7), 54.
- Murphy, D., DeCarli, C., McIntosh, A., Daly, E., Mentis, M., Pietrini, P., Szczepanik, J., Schapiro, M., Grady, C., Horwitz, B., Rapoport, S., 1996. Sex differences in human brain morphology and metabolism: an in vivo quantitative magnetic resonance imaging and positron emission tomography study on the effect of aging. *Arch. Gen. Psychiatry* 53 (7), 585.
- Nathanson, C.A., 1984. Sex differences in mortality. *Annu. Rev. Sociol.* 10, 191–213.
- Newman, J., 1995. Thalamic contributions to attention and consciousness. *Conscious. Cogn.* 4 (2), 172–193.
- Nieuwenhuys, R., Voogd, J., Huijzen, C., 2008. Hypothalamus. In: Nieuwenhuys, R., Voogd, J., Huijzen, C. (Eds.), *The Human Central Nervous System*. Springer, Berlin, pp. 314–320.
- O'Brien, J., Paling, S., Barber, R., Williams, E., Ballard, C., McKeith, I., Gholkar, A., Crum, W., Rossor, M., Fox, N., 2001. Progressive brain atrophy on serial MRI in dementia with Lewy bodies, AD, and vascular dementia. *Neurology* 56 (10), 1386.
- Ohnishi, T., Matsuda, H., Tabira, T., Asada, T., Uno, M., 2001. Changes in brain morphology in Alzheimer disease and normal aging: is Alzheimer's disease an exaggerated aging process? *Am. J. Neuroradiol.* 22 (9), 1680.
- Ott, B., Cahn-Weiner, D., 2001. Gender differences in Alzheimer's disease. *Geriatr. Times* 2 (6).
- Ott, B., Lapane, K., Gambassi, G., 2000. Gender differences in the treatment of behavior problems in Alzheimer's disease. *Neurology* 54 (2), 427.
- Ott, B., Tate, C., Gordon, N., Heindel, W., 1996. Gender differences in the behavioral manifestations of Alzheimer's disease. *J. Am. Geriatr. Soc.* 44 (5), 583.
- Pan, W., 2001. Akaike's information criterion in generalized estimating equations. *Biometrics* 57 (1), 120–125.
- Perrig-Chiello, P., Hutchison, S., 2010. Health and well-being in old age: the pertinence of a gender mainstreaming approach in research. *Gerontology* 56 (2), 208–213.
- Phelps, E., 2004. Human emotion and memory: interactions of the amygdala and hippocampal complex. *Curr. Opin. Neurobiol.* 14 (2), 198–202.
- Piefke, M., Weiss, P., Markowitsch, H., Fink, G., 2005. Gender differences in the functional neuroanatomy of emotional episodic autobiographical memory. *Hum. Brain Mapp.* 24 (4), 313–324.
- Postle, B., D'Esposito, M., 1999. Dissociation of human caudate nucleus activity in spatial and nonspatial working memory: an event-related fMRI study. *Cogn. Brain Res.* 8 (2), 107–115.
- Rabinovici, G., Furst, A., O'Neil, J., Racine, C., Mormino, E., Baker, S., Chetty, S., Patel, P., Pagliaro, T., Klunk, W., Mathis, C.A., Rosen, H.J., Miller, B.L., Jagust, W.J., 2007. 11C-PIB PET imaging in Alzheimer disease and frontotemporal lobar degeneration. *Neurology* 68 (15), 1205.
- Raz, N., Gunning, F.M., Head, D., Dupuis, J.H., McQuain, J., Briggs, S.D., Loken, W.J., Thornton, A.E., Acker, J.D., 1997. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cereb. Cortex* 7, 268–282.
- Rapp, M., Schneider-Beeri, M., Grossman, H., Sano, M., Perl, D., Purohit, D., Gorman, J., Haroutunian, V., 2006. Increased hippocampal plaques and tangles in patients with Alzheimer disease with a lifetime history of major depression. *Arch. Gen. Psychiatry* 63 (2), 161.
- Resnick, S., Goldszal, A., Davatzikos, C., Golski, S., Kraut, M., Metter, E., Bryan, R., Zonderman, A., 2000. One-year age changes in MRI brain volumes in older adults. *Cereb. Cortex* 10 (5), 464.
- Resnick, S., Pham, D., Kraut, M., Zonderman, A., Davatzikos, C., 2003. Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *J. Neurosci.* 23 (8), 3295.
- Ripich, D., Petrill, S., Whitehouse, P., Ziolo, E., 1995. Gender differences in language of AD patients: a longitudinal study. *Neurology* 45 (2), 299.
- Risacher, S., Saykin, A., West, J., Shen, L., Firpi, H., McDonald, B., ADNI, 2009. Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. *Curr. Alzheimer Res.* 6 (4), 347.
- Rocca, W., Hofman, A., Brayne, C., Breteler, M., Clarke, M., Copeland, J., Dartigues, J., Engedal, K., Hagnell, O., Heeren, T., et al., 1991. Frequency and distribution of Alzheimer's disease in Europe: a collaborative study of 1980–1990 prevalence findings. *Ann. Neurol.* 30 (3), 381–390.
- Salat, D., Kaye, J., Janowsky, J., 2001. Selective preservation and degeneration within the prefrontal cortex in aging and Alzheimer disease. *Arch. Neurol.* 58 (9), 1403.
- Scahill, R., Frost, C., Jenkins, R., Whitwell, J., Rossor, M., Fox, N., 2003. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch. Neurol.* 60 (3), 989.
- Schroeter, M., Stein, T., Maslowski, N., Neumann, J., 2009. Neural correlates of Alzheimer's disease and mild cognitive impairment: a systematic and quantitative meta-analysis involving 1351 patients. *Neuroimage* 47 (4), 1196–1206.
- Schuff, N., Tosun, D., Insel, P., Chiang, G., Truran, D., Aisen, P., Jack, C., Weiner, M., 2010. Nonlinear time course of brain volume loss in cognitively normal and impaired elders. *Neurobiol. Aging*.
- Selkoe, D., 2001. Alzheimer's disease: genes, proteins, and therapy. *Physiol. Rev.* 81 (2), 741.

- Shattuck, D., Sandor-Leahy, S., Schaper, K., Rottenberg, D., Leahy, R., 2001. Magnetic resonance image tissue classification using a partial volume model. *Neuroimage* 13 (5), 856–876.
- Sheline, Y., Gado, M., Kraemer, H., 2003. Untreated depression and hippocampal volume loss. *Am. J. Psychiatry* 160, 1516–1518.
- Shen, D., Davatzikos, C., 2003. Very high-resolution morphometry using mass-preserving deformations and HAMMER elastic registration. *Neuroimage* 18 (1), 28–41.
- Shen, D., Davatzikos, C., 2004. Measuring temporal morphological changes robustly in brain MR images via 4-dimensional template warping. *Neuroimage* 21 (4), 1508–1517.
- Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., Baltes, P., 2003. The fate of cognition in very old age: six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychol. Aging* 18, 318–331.
- Sled, J., Zijdenbos, A., Evans, A., 2002. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* 17 (1), 87–97.
- Sluimer, J., Bouwman, F., Vrenken, H., Blankenstein, M., Barkhof, F., van der Flier, W., Scheltens, P., 2010. Whole-brain atrophy rate and CSF biomarker levels in MCI and AD: a longitudinal study. *Neurobiol. Aging* 31 (5), 758–764.
- Sowell, E., Peterson, B., Kan, E., Woods, R., Yoshii, J., Bansal, R., Xu, D., Zhu, H., Thompson, P., Toga, A., 2007. Sex differences in cortical thickness mapped in 176 healthy individuals between 7 and 87 years of age. *Cereb. Cortex* 17, 1550–1560.
- Smith, C., Creed Pettigrew, L., Avison, M., Kirsch, J., Tinkhtman, A., Schmitt, F., Wermeling, D., Wekstein, D., Marchkesberry, W., 1995. Frontal lobe phosphorus metabolism and neuropsychological function in aging and in Alzheimer's disease. *Ann. Neurol.* 38 (2), 194–201.
- Smith, S., 2002. Fast robust automated brain extraction. *Hum. Brain Mapp.* 17 (3), 143–155.
- Stout, J., Bondi, M., Jernigan, T., Archibald, S., Delis, D., Salmon, D., 1999. Regional cerebral volume loss associated with verbal learning and memory in dementia of the Alzheimer type I. *Neuropsychology* 13 (2), 188–197.
- Stuss, D., Guberman, A., Nelson, R., Laroche, S., 1988. The neuropsychology of paramedian thalamic infarction. *Brain Cogn.* 8 (3), 348–378.
- Sullivan, E.V., Marsh, L., Mathalon, D.H., Lim, K.O., Pfefferbaum, A., 1995. Age-related decline in MRI volumes of temporal lobe gray matter but not hippocampus. *Neurobiol. Aging* 16, 591–606.
- Sullivan, E., Rosenbloom, M., Serventi, K., Pfefferbaum, A., 2004. Effects of age and sex on volumes of the thalamus, pons, and cortex. *Neurobiol. Aging* 25 (2), 185–192.
- Teipel, S., Born, C., Ewers, M., Bokde, A., Reiser, M., Moller, H., Hampel, H., 2007. Multivariate deformation-based analysis of brain atrophy to predict Alzheimer's disease in mild cognitive impairment. *Neuroimage* 38 (1), 13–24.
- Thambisetty, M., Wan, J., Carass, A., An, Y., Prince, J., Resnick, S., 2010. Longitudinal changes in cortical thickness associated with normal aging. *Neuroimage* 52 (4), 1215–1223.
- Tranel, D., Damasio, H., Damasio, A., 1997. A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia* 35 (10), 1319–1327.
- Tranel, D., Hyman, B., 1990. Neuropsychological correlates of bilateral amygdala damage. *Arch. Neurol.* 47 (3), 349.
- Valera, E., Brown, A., Biederman, J., Faraone, S., Makris, N., Monuteaux, M., Whitfield-Gabrieli, S., Vitulano, M., Schiller, M., Seidman, L., 2010. Sex differences in the functional neuroanatomy of working memory in adults with ADHD. *Am. J. Psychiatry* 167 (1), 86.
- Van der Werf, Y., Witter, M., Uylings, H., Jolles, J., 2000. Neuropsychology of infarctions in the thalamus: a review. *Neuropsychologia* 38, 613–627.
- Vina, J., Lloret, A., 2010. Why women have more Alzheimer's disease than men: gender and mitochondrial toxicity of amyloid- β peptide. *J. Alzheimers Dis.* 20, 527–533.
- Walhovd, K., Westlye, L., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Salat, D., Greve, D., Fischl, B., Dale, A., Fjell, A., 2009. Consistent neuroanatomical age-related volume differences across multiple samples. *Neurobiol. Aging*.
- Wiederholt, W., Cahn, D., Butters, N., Salmon, D., Kritz-Silverstein, D., Barrett-Connor, E., 1993. Effects of age, gender and education on selected neuropsychological tests in an elderly community cohort. *J. Am. Geriatr. Soc.* 41, 639–647.
- Witte, A., Savli, M., Holik, A., Kasper, S., Lanzenberger, R., 2010. Regional sex differences in grey matter volume are associated with sex hormones in the young adult human brain. *Neuroimage* 49 (2), 1205–1212.
- Xu, J., Kobayashi, S., Yamaguchi, S., Iijima, K., Okada, K., Yamashita, K., 2000. Gender effects on age-related changes in brain structure. *Am. J. Neuroradiol.* 21 (1), 112.
- Yotsutsuji, T., Saitoh, O., Suzuki, M., Hagino, H., Mori, K., Takahashi, T., Kurokawa, K., Matsui, M., Seto, H., Kurachi, M., 2003. Quantification of lateral ventricular subdivisions in schizophrenia by high resolution three-dimensional magnetic resonance imaging. *Psychiatry research. Neuroimaging* 122 (1), 1–12.
- Zhang, Y., Brady, M., Smith, S., 2002. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20 (1), 45–57.
- Zhu, H., Li, Y., Ibrahim, J., Lin, W., Shen, D., 2009. MARM: Multiscale Adaptive Regression Models for neuroimaging data. *Information Processing in Medical Imaging. Springer*, pp. 314–325.