Original Article

Journal of INTERNAL MEDICINE

Click here to view the Editorial Comment by P. Scheltens

doi: 10.1111/joim.12148

Influence of age, disease onset and *ApoE4* on visual medial temporal lobe atrophy cut-offs

J. B. Pereira[1], L. Cavallin[2,3], G. Spulber[1], C. Aguilar[1], P. Mecocci[4], B. Vellas[5], M. Tsolaki[6], I. Kłoszewska[7], H. Soininen[8], C. Spenger[2], D. Aarsland[1], S. Lovestone[9,10], A. Simmons[9,10,11], L.-O. Wahlund[1], E. Westman[1] & for the AddNeuroMed consortium and for the Alzheimer's Disease Neuroimaging Initiative*

From the [1]Department of Neurobiology, Care Sciences and Society, Karolinska Institutet; [2]Division of Medical Imaging and Technology, Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm; [3]Department of Radiology, Karolinska University Hospital in Huddinge, Huddinge, Sweden; [4]Institute of Gerontology and Geriatrics, University of Perugia, Perugia, Italy; [5]INSERM U 558, University of Toulouse, Toulouse, France; [6]Aristotle University of Thessaloniki, Thessaloniki, Greece; [7]Medical University of Lodz, Lodz, Poland; [8]University of Eastern Finland, University Hospital of Kuopio, Kuopio, Finland; [9]Institute of Psychiatry, King's College London; [10]NIHR Biomedical Research Centre for Mental Health; and [11]NIHR Biomedical Research Unit for Dementia, London, UK

Abstract. Pereira JB, Cavallin L, Spulber G, Aguilar C, Mecocci P, Vellas B, Tsolaki M, Kłoszewska I, Soininen H, Spenger C, Aarsland D, Lovestone S, Simmons A, Wahlund L-O, Westman E (Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Novum 5th floor, SE-141 86 Stockholm, Sweden). Influence of age, disease onset and *ApoE4* on visual medial temporal lobe atrophy cut-offs. *J Intern Med* 2014; **275**: 317–330.

Background. Visual assessment of medial temporal lobe atrophy (MTA; range 0–4, from no atrophy to increasing atrophy of the choroid fissure, temporal horns and hippocampus) is a sensitive radiological marker of Alzheimer's disease (AD). One of the critical elements for visual MTA assessment is the cut-off score that determines deviation from normality.

Methods. In this study, we assessed the sensitivity and specificity of different MTA cut-off scores to classify control subjects, individuals with mild cognitive impairment (MCI) and AD patients from two large independent cohorts, AddNeuroMed and Alzheimer's Disease Neuroimaging Initiative. Of note, we evaluated the effects of clinical, demographic and genetic variables on the classification performance according to the different cut-offs.

Results. A cut-off of ≥1.5 based on the mean MTA scores of both hemispheres showed higher sensitivity in classifying patients with AD (84.5%) and MCI subjects (75.8%) who converted to dementia compared to an age-dependent cut-off. The age-dependent cut-off showed higher specificity or ability to correctly identify control subjects (83.2%) and those with MCI who remained stable (65.5%). Increasing age, early-onset disease and absence of the *ApoE* ε4 allele had a stronger influence on classifications using the ≥1.5 cut-off. Above 75 years of age, an alternative cut-off of ≥2.0 should be applied to achieve a classification accuracy for both patients with AD and control subjects that is clinically useful.

Conclusion. Clinical, demographic and genetic variables can influence the classification of MTA cut-off scores, leading to misdiagnosis in some cases. These variables, in addition to the differential sensitivity and specificity of each cut-off, should be carefully considered when performing visual MTA assessment.

Keywords: AddNeuroMed, Alzheimer's disease, Alzheimer's Disease Neuroimaging Initiative, magnetic resonance imaging, medial temporal lobe atrophy.

Introduction

Nondemented elderly individuals with persistent memory problems may be in an intermediate stage between normal ageing and Alzheimer's disease (AD), known as amnestic mild cognitive impairment (MCI) [1]. However, the uncertainty of progression to dementia amongst those with MCI and the

*Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data, but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://supp.apa.org/psycarticles/supplemental/a0034057/ADNI_Acknowledgement_List.pdf.

increasing prevalence of memory deficits in the ageing population [2] suggest an urgent need for supportive measures that can help the clinician to monitor these individuals [3].

Magnetic resonance imaging (MRI) is a commonly available tool with predictive value for dementia [4]. Currently, MRI scans are used in the routine workup of memory disorders for excluding non-AD pathologies such as hydrocephalus, brain tumours, haematomas and strokes as the cause of the cognitive syndrome. However, these scans can also be used to support evaluation of the presence and severity of AD and serve as a predementia marker in the clinical setting [5].

Quantitative MRI studies have shown a pattern of progression of brain atrophy in AD affecting the medial temporal lobe, especially the hippocampus [6–8]. This pattern not only reliably discriminates healthy ageing from AD and other dementia disorders associated with medial temporal lobe atrophy (MTA) [9], but also predicts progression to dementia in individuals with MCI [10–13]. However, despite providing important diagnostic information about AD and other dementias, the application of quantitative MRI analysis in clinical practice is not routine. This is partly because quantitative MRI measures are not widely available, require previous training and specialized software and are therefore not readily usable by clinicians. In 1992, Scheltens *et al.* [14] developed a visual rating scale to grade the severity of MTA that is quick and easy to use and can be immediately implemented in a clinical setting. This scale has a range of 0 (no atrophy) to 4 (maximum atrophy) and is based on the evaluation of a number of anatomical features that include the hippocampus, parahippocampal gyrus, entorhinal cortex and surrounding cerebrospinal fluid (CSF) spaces such as the temporal horn and choroid fissure.

Previous studies have shown that the MTA scale is useful for identifying prodromal stages of AD [15–17] and distinguishing patients with AD from those with other dementias [18]. In addition, MTA assessment was found to be a highly sensitive marker of memory impairment [19] and AD-related neuropathology [18]. Although MTA is a sensitive early marker of AD, it is also commonly found in healthy elderly subjects, particularly in those of advanced age [20], as well as in patients with frontotemporal dementia [21], dementia with Lewy bodies [22] and Parkinson's disease with dementia [23]. Moreover, the absence of significant MTA does not exclude the diagnosis of AD, especially in patients with early-onset disease [24, 25]. In these cases, the presence of MTA is often not clear and a different pattern of brain atrophy involving the inferior parietal cortex (angular and supramarginal gyri) as well as the precuneus is observed instead [26–28]. In addition to these issues, certain factors might also influence the extent and predisposition to increased MTA such as carrying the *ApoE* allele ε4 [29], a known risk factor for developing AD [30].

In view of the impact of demographic, clinical and genetic factors on MTA in the prodromal and clinical stages of AD, one of the critical elements for the visual MTA assessment is the cut-off score that determines deviation from normality [31, 32]. Scheltens *et al.* [14] proposed that an MTA score ≥2 is abnormal below the age of 75, whilst above 75 years of age a score ≥3 would be required to identify abnormality in either hemisphere. However, this age-dependent cut-off often misclassifies many patients with AD, leading to a low sensitivity [33]. In line with this, several studies have used different cut-offs lower than the score originally proposed by Scheltens *et al.* [14] that are based on the average of the MTA scores of both hemispheres. For example, a cut-off score of ≥1.5 has been used [34, 35]; this cut-off requires a minimum MTA score of 1 in one hemisphere characterized by increased width of the choroid fissure and an MTA score of 2 in the other hemisphere characterized by increases in the width of the choroid fissure and temporal horns as well as hippocampal thinning. Although higher sensitivity might be accompanied by lower specificity, few studies have addressed this important issue.

The main goal of the current study was to assess the sensitivity and specificity of two different MTA cut-offs (≥1.5 and age dependent) to discriminate and predict conversion to AD in two large multicentre cohorts: AddNeuroMed and the Alzheimer's Disease Neuroimaging Initiative (ADNI). In a previous study, we assessed visual MTA in part of the AddNeuroMed cohort and found that an age-based cut-off distinguished patients with AD from control subjects with 81% accuracy, in addition to predicting future conversion to AD in 68% of the MCI cases [36]. In another study [37], visual MTA was also assessed in the ADNI cohort, with a strong association being found between MTA ratings and conversion to AD. However, in these studies, no comparisons were made between different cut-offs

or between the two large multicentre cohorts. In addition, the important influence of the characteristics of individual patients on MTA ratings was not investigated in either of these previous studies [36, 37]. Hence, in the current study, by comparing the classification accuracies of different MTA cut-offs in the largest number of patients ever assessed to date with the visual MTA scale, we aimed to provide clinicians with information on the best way to define the presence of medial temporal lobe abnormalities in their patients. Moreover, we investigated the impact of demographic, clinical and genetic factors on both cut-offs (≥1.5 and age dependent) that might be responsible for misdiagnosis and in this way show in which cases visual MTA assessment may be inaccurate.

Methods

Subjects

A total of 1147 individuals (345 control subjects, 480 subjects with MCI and clinical follow-up at 1 year, 322 patients with AD) were included in this study. Subjects with MCI and AD patients were recruited through local hospital and memory clinics at the ADNI and AddNeuroMed participating sites; control subjects were recruited through advertisements in newspapers for the ADNI study and from unrelated members of the patients' families, caregivers' relatives, social centres for the elderly or GP surgeries for the AddNeuroMed study.

Data used in the present study were obtained from the AddNeuroMed study and the ADNI database (www.adni-info.org). The AddNeuroMed study is part of the InnoMed European Union FP6 programme and was designed to develop and validate novel surrogate markers in AD [38, 39]. 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 nonprofit organizations, as a $60 million, 5-year public–private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) 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 to lessen the time and cost of clinical trials. The principal investigator of this initiative is Michael W. Weiner, MD, 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 USA and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date, these three protocols have recruited over 1500 adults, aged 55–90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI and people with early AD. The follow-up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

For both the AddNeuroMed study and the ADNI, the inclusion criteria for the AD group were as follows: (i) NINCDS-ADRDA [40] and DSM-IV criteria for probable AD, (ii) Mini-Mental State Examination (MMSE) [41] score ranging from 12 to 28, (iii) age 65 years or above and (iv) total Clinical Dementia Rating (CDR) score of 0.5 or above. Exclusion criteria were as follows: (i) significant neurological or psychiatric illness other than AD and (ii) significant unstable systemic illness or organ failure.

The inclusion criteria for the MCI group were as follows: (i) MMSE score between 24 and 30 [41], (ii) memory complaint reported by the patient, family member or physician, (iii) normal activities of daily living, (iv) CDR memory score of 0.5 or 1 (total CDR = 0.5), (v) memory loss measured by the Wechsler Memory Scale Logical Memory II for the ADNI cohort only (education-adjusted scores: ≤8 for 16 or more years of education, ≤4 for 8–15 years of education, ≤2 for 0–7 years of education), (vi) Geriatric Depression Scale (GDS) score of ≤5, (vii) age ≥65 years, (viii) stable medication and (ix) good general health. Exclusion criteria were as follows: (i) meeting the DSM-IV and NINCDS-ADRDA criteria [40] for AD, (ii) significant neurological or psychiatric illness other than AD and (iii) significant unstable systemic illness or organ failure.

The inclusion criteria for control subjects were as follows: (i) MMSE score between 24 and 30 [41], (ii) GDS score < 5, (iii) age ≥ 65 years, (iv) stable medication and (v) good general health. Exclusion

criteria were as follows: (i) DSM-IV or NINCDS-ADRDA criteria [40] for dementia, (ii) significant neurological or psychiatric illness other than AD and (iii) significant unstable systemic illness or organ failure.

The CERAD cognitive battery was used to assess control subjects and those with MCI, whereas for patients with AD we used the Alzheimer's disease Assessment Scale-cognitive subscale (ADAS-Cog), which is specifically designed for AD trials [42]. Both the ADAS-Cog and the CERAD battery use the same 10-word recall task, although the scoring in the two tests is in the opposite direction. The mean number of words that were not recalled in the word list of the CERAD immediate recall task was calculated and the variable obtained was termed ADAS1, corresponding to the first subtest of ADAS-Cog. This calculation was performed to provide comparable measures for the ADNI and AddNeuroMed cohorts. For ADNI, *ApoE* genotyping was determined using EDTA-anticoagulated blood samples and TaqMan assays, as described elsewhere [43]. For AddNeuroMed, *ApoE* genotype was determined from leucocytes and analysed by polymerase chain reaction, HhaI digestion and polyacrylamide gel electrophoresis [44]. All clinical diagnoses, for both cohorts, were made without the use of MRI scans.

Magnetic resonance imaging

Magnetic resonance imaging data acquisition in the AddNeuroMed study was designed to be compatible with that in ADNI [45]. For both studies, the imaging protocol involved a high-resolution sagittal 3D T1-weighted magnetization prepared rapid gradient echo (MPRAGE) volume (voxel size $1.1 \times 1.1 \times 1.2$ mm^3) and axial proton density-/T2-weighted fast spin echo images. The MPRAGE volume was obtained using a customized pulse sequence specifically designed for the ADNI study to ensure compatibility across scanners [44]. Full brain and skull coverage was required, and detailed quality control was carried out on all MRI data according to the AddNeuroMed quality control procedure [46, 47].

Visual assessment of MTA

In both cohorts, 3D T1-weighted images from all participants were reoriented to an oblique coronal orientation perpendicular to the anterior commissure–posterior commissure line, appropriate for visual and quantitative MRI assessment. For each subject, MTA was rated on a single magnetic resonance slice posterior to the amygdala and the mammillary bodies, positioned in such a way that the hippocampus, cerebral peduncles and pons were all visible. Then, we used the rating scale proposed by Scheltens *et al.*, which is based on a visual estimation of the volume of the medial temporal lobe including the hippocampus, parahippocampal gyrus, entorhinal cortex, temporal horns and choroid fissure. According to this scale, scores range from 0 (no atrophy) to 4 (end-stage atrophy) based on the width of the choroid fissure, width of the temporal horn and hippocampal thickness. The right and left sides of the medial temporal lobe are rated separately.

Deviation from normality was determined by using two different and independent cut-off values. First, regarding the age-dependent cut-off, an MTA score of 2 or more was considered abnormal for subjects <75 years, whereas a score of 3 or more was considered abnormal for subjects >75 years. According to this age-dependent scoring method, it is sufficient to have an abnormal score in one hemisphere for the subject to be classified as having MTA. The second cut-off was calculated based on the average of the MTA scores of both hemispheres, with a resulting score equal to or greater than 1.5 being considered abnormal.

In the current study, the MTA rater (LC) was blind to gender, age and diagnosis. Intrarater reliability of MTA assessment was tested in 100 randomly selected subjects by repeated assessment at an interval of 1 week and was found to be 0.93 and 0.94 (weighted kappa) for the left and right hemispheres, respectively. Visual MTA ratings have been previously performed by LC in other large samples such as the Swedish National Study on Ageing and Care in Kungshomen including 544 elderly individuals. A highly significant correlation was found between the MTA score and manual delineation of hippocampal volumes by another experienced radiologist [48].

Statistical analyses

Statistical analyses were carried out using SPSS version 16.0 (SPSS Inc., Chicago, IL, USA). To assess the ability of the visual MTA scale to discriminate patients with AD from control subjects as well as predicting conversion from MCI to AD at 1-year follow-up, sensitivity, specificity and

accuracy (and the corresponding confidence intervals) were calculated based on the MTA score of ≥1.5 and age-dependent cut-offs. These calculations were made in both cohorts, separately and combined. To test whether performance in classifying the clinical groups was significantly different between the two MTA cut-offs, the areas under the curve (AUCs) were compared using Med-Calc version 12.2.1 (MedCalc Software, Mariakerke, Belgium). To compare separately the sensitivities and specificities between the two types of cut-off, McNemar's chi-square tests were applied using the following formula: $\chi^2 = (|c-b|-1)^2/(c+b)$ [49], where b represents the true-positive values according to the ≥1.5 cut-off and false-negative values according to the age cut-off, and c represents the true-positive values according to the age cut-off and the false-negative values according to the ≥1.5 cut-off. Differences were considered statistically significant at $P < 0.05$, if $\chi^2 > 3.841$ (one degree of freedom).

In order to evaluate the impact of clinical and demographic characteristics on the ≥1.5 and age-dependent scoring methods, comparisons between correctly and incorrectly classified subjects were made. One-way analysis of variance (ANOVA) was used for continuous clinical and demographic variables, a chi-squared test for binary variables such as gender or *ApoE* ε4 prevalence, and a Kruskal–Wallis test for ordinal variables such as CDR scores. To further explore the influence of clinical variables on MTA, nonparametric comparisons of MTA cut-offs using the Mann–Whitney test were also made between (i) patients with early-onset (before 65 years of age) and late-onset (at or after 65 years of age) AD, (ii) AD patients carrying the *ApoE* ε4 allele and noncarriers and (iii) AD *ApoE* ε4 carriers and noncarriers with early- and late-onset disease.

The influence of age on MTA scores was evaluated by classifying control subjects and patients with AD into two age groups (50–74 and 75–90 years) within the entire cohort. The ability of the ≥1.5 and age-dependent MTA cut-offs to correctly classify patients with AD and healthy control subjects within these different age ranges was calculated. In addition, we assessed the classification performance of another cut-off (score of ≥2.0) based on the mean of both hemispheres to determine whether a higher threshold could improve the discrimination between older patients with AD and control subjects [20].

Results

Baseline characteristics of individuals in the control, MCI and AD groups in the ADNI, AddNeuroMed and combined cohorts are shown in Table 1. A total of 70 subjects with MCI from the ADNI cohort and 25 from the AddNeuroMed study were diagnosed with AD after 12 months, whilst a total of 284 and 101 MCI subjects remained stable from the two cohorts, respectively (Table 1).

Classification accuracy of the different MTA cut-offs

Figure 1 shows the MTA scores in two cases using the ≥1.5 and age-dependent cut-offs. With regard to the different cut-offs, an average MTA score of ≤1 or a maximum score of 1 in both hemispheres is always considered normal, whereas an average MTA score of 2 or a minimum score of 3 in one hemisphere is always considered abnormal. The two cut-offs differ in terms of the classification of subjects with an MTA score of 2 in at least one hemisphere; according to the ≥1.5 cut-off, subjects will always be classified as abnormal only if they have an average MTA score of ≥1.5, whereas the age-dependent cut-off takes into account the age and assigns an abnormal score only if an individual is younger than 75 years.

Regarding the discrimination of patients with AD from control subjects, we observed that a cut-off score of ≥1.5 resulted in a classification accuracy of 74.4% in the ADNI, 79% in the AddNeuroMed study and 76% in the combined cohort (Table 2). The accuracy of the age-dependent cut-off was similar to that of the ≥1.5 cut-off, but with lower sensitivity and higher specificity, suggesting that this was a more conservative approach (Table 2). Although there were no significant differences between the AUC values of the two cut-offs (Supplementary Table 1), the ≥1.5 cut-off showed a significantly higher sensitivity in identifying patients with AD in the ADNI ($\chi^2 = 26.036$, $P < 0.05$), AddNeuroMed study ($\chi^2 = 17.053$, $P < 0.05$) and the combined cohort ($\chi^2 = 45.020$, $P < 0.05$). By contrast, the age-dependent cut-off showed a significantly higher specificity in these cohorts ($\chi^2 = 38.205$, $P < 0.05$; $\chi^2 = 8.100$, $P < 0.05$; and $\chi^2 = 48.167$, $P < 0.05$, respectively), compared to the ≥1.5 cut-off.

Predicting conversion from MCI to AD

Of 95 MCI subjects in the combined cohort who converted to AD at the 1-year follow-up, 72 were

Table 1 *Baseline subject characteristics*

	ADNI cohort				AddNeuroMed cohort				Combined cohort			
	AD (n = 199)	MCI-c (n = 70)	MCI-s (n = 284)	CTR (n = 230)	AD (n = 123)	MCI-c (n = 25)	MCI-s (n = 101)	CTR (n = 115)	AD (n = 322)	MCI-c (n = 95)	MCI-s (n = 385)	CTR (n = 345)
Age	75.5 (7.7)	74.5 (6.6)	75.2 (7.3)	75.9 (5.0)	75.7 (6.1)	73.5 (6.3)	74.8 (5.5)	72.9 (6.7)	75.6 (7.1)	74.3 (6.5)	75.1 (6.9)	74.9 (5.8)
Sex (M/F)	103/96	42/28	184/100	118/112	43/80	14/11	49/52	51/64	146/176	56/39	233/152	169/176
Education	14.7 (3.2)	15.4 (2.8)	15.8 (3.0)	16.0 (2.9)	8.0 (4.0)	9.2 (4.2)	8.8 (4.3)	10.8 (4.8)	12.2 (4.8)	13.7 (4.2)	14.0 (4.6)	14.3 (4.4)
CDR	0.7 (0.3)	0.5	0.5	0	1.2 (0.5)	0.5	0.5	0	0.9 (0.4)	0.5	0.5	0
MMSE	22.0 (5.7)	26.5 (1.8)	27.1 (1.7)	28.9 (2.9)	20.8 (4.8)	26.6 (1.9)	27.2 (1.6)	29.0 (1.2)	21.5 (5.4)	26.5 (1.8)	27.1 (1.7)	28.9 (2.5)
ADAS1	6.1 (1.4)	5.4 (1.3)	4.4 (1.3)	2.8 (1.2)	6.6 (1.5)	5.3 (1.4)	5.3 (1.2)	3.5 (1.5)	6.3 (1.5)	5.3 (1.3)	4.6 (1.4)	3.1 (1.3)
ApoE ε4 (%)	66.3	61.4	53.2	26.5	51.2	64.0	25.7	29.6	60.6	62.1	46	27.5

AD, Alzheimer's disease; CTR, control; MCI-c, subjects with conversion after 1 year of mild cognitive impairment to dementia; MCI-s, subjects with mild cognitive impairment that remained stable after 1 year; MMSE, Mini-Mental State Examination; CDR, Clinical Dementia Rating; ADAS1, Alzheimer's Disease Assessment Scale memory subtest; *ApoE* ε4, apolipoprotein E allele ε4; ADNI, Alzheimer's Disease Neuroimaging Initiative. All values are presented as means followed by standard deviation, except for sex and *ApoE* ε4 prevalence.

Fig. 1 *Differences between the ≥1.5 and age-dependent cut-offs of medial temporal lobe atrophy (MTA). Using both cut-offs, a score lower than or equal to 1 is always considered normal, whilst a score greater than 2 is always considered abnormal. The difference between cut-offs resides in the classification of subjects with MTA scores >1 and ≤2. Whilst the ≥1.5 cut-off, based on the mean score of both hemispheres, always classifies subjects between the above scores as abnormal, the age-dependent cut-off only classifies as abnormal subjects below 75 years of age. In the combined cohort, 15.5% of patients with Alzheimer's disease, 40.6% of mild cognitive impairment subjects and 68.1% of control subjects were classified as normal (MTA ≤ 1); 52.2%, 20.9% and 5.8%, respectively, were classified as abnormal (MTA > 2); and 32.3%, 38.5% and 26.1%, respectively, had a score between ≤1 and >2.*

identified as more AD-like (75.8%) using the ≥1.5 cut-off, whilst the age-dependent cut-off only correctly classified 57 (60%) MCI converters (Table 3). Although the AUC values did not show significant differences (Table S1), the ≥1.5 cut-off had significantly higher sensitivity in the ADNI ($\chi^2 = 38.205$, $P < 0.05$) and the combined cohorts ($\chi^2 = 13.067$, $P < 0.05$), but not in the AddNeuroMed cohort ($\chi^2 = 0.5$). On the other hand, the age-dependent cut-off showed significantly higher specificity in all previous cohorts ($\chi^2 = 60.063$, $P < 0.05$; $\chi^2 = 80.105$, $P < 0.05$; and $\chi^2 = 16.409$, $P < 0.05$, respectively), compared to the ≥1.5 cut-off.

Characteristics of correctly versus incorrectly classified subjects

Using the ≥1.5 MTA cut-off, patients with AD who were incorrectly classified as more control-like were younger ($P < 0.001$), had better CDR and ADAS1

Table 2 *Sensitivity, specificity and accuracy for the AD versus control classification by the different MTA cut-offs*

	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)
MTA ≥1.5 cut-off			
ADNI	82.9 (77.1–87.5)	67 (60.6–72.7)	74.4 (70–78.3)
AddNeuroMed	87 (79.9–91.8)	70.4 (61.5–78)	79 (73.4–83.7)
Combined	84.5 (80.1–88)	68.1 (63–72.8)	76 (72.6–79.1)
MTA age cut-off			
ADNI	68.8 (62.1–74.9)	85.2 (80.1–89.2)	77.6 (73.4–81.3)
AddNeuroMed	71.5 (63–78.8)	79.1 (70.8–85.6)	75.2 (69.4–80.3)
Combined	69.9 (64.7–74.6)	83.2 (78.9–86.8)	76.8 (73.4–79.8)

ADNI, Alzheimer's Disease Neuroimaging Initiative; CI, confidence interval; AD, Alzheimer's disease; MTA, medial temporal atrophy.

Table 3 *Prediction of conversion to dementia in individuals with MCI according to the different MTA cut-offs*

	n	AD-like, *n* (%)	CTR-like, *n* (%)
MTA ≥1.5 cut-off			
ADNI converters	70	52 (74.3)	18 (25.7)
AddNeuroMed converters	25	20 (80)	5 (20)
Combined converters	95	72 (75.8)	23 (24.2)
ADNI nonconverters	284	166 (58.5)	118 (41.6)
AddNeuroMed nonconverters	101	51 (50.5)	50 (49.5)
Combined nonconverters	385	217 (56.4)	168 (43.6)
MTA age-dependent cut-off			
ADNI converters	70	39 (55.7)	31 (44.3)
AddNeuroMed converters	25	18 (72)	7 (28)
Combined converters	95	57 (60)	38 (40)
ADNI nonconverters	284	102 (35.9)	182 (64.1)
AddNeuroMed nonconverters	101	31 (30.7)	70 (69.3)
Combined nonconverters	385	133 (34.5)	252 (65.5)

ADNI, Alzheimer's Disease Neuroimaging Initiative; AD, Alzheimer's disease; CTR, control; MCI, mild cognitive impairment; MTA, medial temporal atrophy.

scores ($P < 0.009$ and $P < 0.004$, respectively), a lower prevalence of the ε4 allele ($P < 0.039$) and a shorter disease duration ($P < 0.041$), were more likely to be male ($P < 0.007$) and were younger at disease onset ($P < 0.006$). On the other hand, control subjects who were classified as having a more AD-like pattern were significantly older ($P < 0.001$) but also more likely to be male ($P < 0.019$). Finally, in MCI nonconverters, those with a more AD-like pattern were older ($P < 0.001$), had worse MMSE and ADAS1 scores ($P < 0.007$ and $P < 0.027$, respectively), showed a higher prevalence of the ε4 allele ($P < 0.011$) and were more likely to be male ($P < 0.025$) according to the ≥1.5 cut-off. Using the age-dependent cut-off, we found that patients with AD incorrectly classified as control-like showed better ADAS1 scores ($P < 0.001$), whilst MCI nonconverters incorrectly classified as more AD-like had a higher prevalence of the ε4 allele ($P < 0.022$; Table 4).

Influence of age at disease onset and ApoE ε4 allele

To evaluate the influence of age at disease onset on the MTA cut-offs, patients with AD from the combined cohort were classified into early-onset (<65 years old; $n = 52$) and late-onset (≥65 years old; $n = 268$) groups. A lower degree of MTA was found in patients with early-onset disease, compared to those with late-onset disease, using the ≥1.5 ($P < 0.001$), but not the age-dependent cut-off ($P < 0.644$). Patients with AD were also grouped based on whether ($n = 195$) or not they carried ($n = 119$) the *ApoE* ε4 allele. Similar to the above group comparison, we found a higher degree of MTA in carriers compared to noncarriers using the ≥1.5 ($P < 0.040$), but not the age-dependent cut-off ($P < 0.392$).

To assess whether age at disease onset and *ApoE* ε4 allele had a combined effect on temporal atrophy, patients with AD were divided into four groups: early-onset noncarriers ($n = 17$), early-onset

Table 4 *Characteristics of the clinical groups based on the classifications of the ≥1.5 and age-dependent MTA cut-offs for the combined cohort*

≥1.5 cut-off	Age	Sex (M/F)	Education	MMSE	CDR	ADAS1	*ApoE* e4	Disease onset	Disease duration
AD									
AD-like	76.2 (6.6)	132/140	12.1 (4.7)	20.9 (7.5)	0.9 (0.4)	6.4 (1.4)	62.9%	72.4 (7.1)	3.8 (2.6)
CTR-like	72.2 (9.1)	14/36	12.5 (5.2)	21.7 (4.9)	0.8 (0.3)	5.8 (1.4)	48%	69.2 (9.4)	2.0 (1.9)
P value	<0.001	0.007	0.597	0.367	0.009	0.004	0.039	0.006	0.041
Control									
AD-like	77.2 (5.7)	64/46	14.2 (4.4)	28.9 (1.2)	0	3.2 (1.4)	30%	–	–
CTR-like	73.8 (5.7)	105/130	14.5 (4.3)	28.9 (2.9)	0	3.0 (1.3)	26.4%	–	–
P value	<0.001	0.019	0.569	0.997	–	0.080	0.510	–	–
MCI-c									
AD-like	74.9 (6.4)	44/28	13.9 (4.2)	26.5 (1.9)	0.5	5.4 (1.4)	65.3%	–	–
CTR-like	72.3 (6.8)	12/11	13.1 (4.5)	26.6 (1.6)	0.5	5.1 (1.3)	52.2%	–	–
P value	0.097	0.448	0.408	0.957	–	0.299	0.321	–	–
MCI-s									
AD-like	76.8 (5.9)	142/75	13.9 (4.7)	26.9 (1.7)	0.5	4.8 (1.3)	51.6%	–	–
CTR-like	72.8 (7.4)	91/77	14.2 (4.4)	27.4 (1.7)	0.5	4.4 (1.5)	38.7%	–	–
P value	<0.001	0.025	0.561	0.007	–	0.027	0.011	–	–

Age-dependent cut-off	Age	Sex (M/F)	Education	MMSE	CDR	ADAS1	*ApoE* e4	Disease onset	Disease duration
AD									
AD-like	75.4 (6.7)	110/115	12.2 (4.7)	21.5 (4.9)	1.0 (0.5)	6.5 (1.4)	63.6%	71.5 (7.1)	3.9 (2.5)
CTR-like	76.1 (8.0)	36/61	12.1 (5.0)	21.5 (6.5)	0.8 (0.3)	5.9 (1.4)	58.5%	72.8 (8.5)	3.3 (2.5)
P value	0.381	0.051	0.911	0.941	0.06	<0.001	0.391	0.159	0.059
Control									
AD-like	74.3 (5.3)	32/26	13.3 (4.2)	28.8 (1.3)	0	3.3 (1.5)	33.3%	–	–
CTR-like	75.0 (5.9)	137/150	14.5 (4.4)	28.9 (2.6)	0	3.0 (1.3)	26.9%	–	–
P value	0.410	0.301	0.059	0.757	–	0.107	0.320	–	–
MCI-c									
AD-like	73.9 (6.7)	37/20	13.8 (4.4)	26.8 (2.0)	0.5	5.4 (1.3)	67.9%	–	–
CTR-like	74.8 (6.3)	19/19	13.6 (4.0)	26.2 (1.5)	0.5	5.3 (1.4)	56.8%	–	–
P value	0.509	0.148	0.860	0.175	–	0.605	0.277	–	–
MCI-s									
AD-like	74.4 (6.0)	88/45	14.2 (4.7)	27.0 (1.7)	0.5	4.7 (1.2)	55.3%	–	–
CTR-like	75.4 (7.3)	145/107	13.9 (4.5)	27.2 (1.7)	0.5	4.6 (1.4)	43.0%	–	–
P value	0.181	0.100	0.506	0.175	–	0.269	0.022	–	–

AD, Alzheimer's disease; CTR, control; MCI-c, subjects with conversion after 1 year of mild cognitive impairment to dementia; MCI-s, subjects with mild cognitive impairment that remained stable after 1 year; MMSE, Mini-Mental State Examination; CDR, Clinical Dementia Rating; ADAS1, Alzheimer's Disease Assessment Scale memory subtest.
All values are presented as means followed by standard deviation, except for sex and *ApoE* ε4 prevalence.

carriers (n = 33), late-onset noncarriers (n = 101) and late-onset carriers (n = 162). These analyses showed significant differences between early-onset noncarriers and ε4 carriers using both the ≥1.5 ($P < 0.012$) and age-dependent ($P < 0.024$) cut-offs. By contrast, no differences were found between late-onset noncarriers and ε4 carriers using either cut-off (≥1.5, $P < 0.267$; age-dependent, $P < 0.990$). In

noncarriers, there were significant differences between those with early-onset and those with late-onset disease using the ≥1.5 cut-off ($P < 0.001$). By contrast, in patients who carried the ε4 allele, there were no significant differences in MTA between those with early- and late-onset disease using either cut-off (≥1.5, $P < 0.261$; age-dependent, $P < 0.329$).

Classification performance of MTA cut-offs in different age groups

To assess the effect of age on MTA cut-offs, patients with AD and control subjects were divided into younger and older age groups in the combined cohort. Between 50 and 74 years of age, the performance of the two cut-offs in classifying patients with AD and control subjects was almost identical. However, from 75 to 90 years of age, the ≥1.5 cut-off showed high sensitivity and low specificity, whereas the age cut-off showed the opposite pattern (Table 5). By applying an alternative cut-off of ≥2.0 based on the average of both hemispheres, we found that almost 82% of patients with AD and 75% of control subjects in the older age group were correctly identified, suggesting that this was the best classification approach for subjects between 75 and 90 years of age.

Discussion

Many elderly people report a decline in memory compared to earlier in their lives [50], making it difficult to detect the first signs of AD [4]. The presence of MCI is common in individuals older than 65 years of age [2], suggesting an urgent need for measures that can assist the clinician in predicting which individuals are most likely to progress to AD. In this study, we have shown that a cut-off score of ≥1.5 in visual MTA achieved a higher sensitivity in correctly classifying patients with AD and predicting conversion from MCI to AD compared to an age-based cut-off. On the other hand, the age-based cut-off showed a significantly higher specificity or ability to correctly identify healthy control and MCI subjects who remained stable. The classification of both cut-offs was influenced by several variables, such as increasing age, time of disease onset and presence of the *ApoE* ε4 allele, leading to misdiagnosis. These variables, in addition to the differential sensitivity and specificity of each type of cut-off, should be carefully considered when performing visual MTA assessment.

In AD, involvement of the hippocampus and surrounding medial temporal lobe is one of the most characteristic histopathological events [51]. For this reason, new diagnostic criteria that capture the earliest signs of AD include structural neuroimaging biomarkers that are focused on the assessment of MTA [52, 53]. The inclusion of visual MTA measures in clinical examinations could indeed improve the sensitivity and specificity of AD diagnosis. In the present study, we observed that an MTA score of ≥1.5 correctly identified AD patients with 82.9% and 87% sensitivity in the ADNI and AddNeuroMed cohorts, respectively. This is similar to previous findings using automated MRI measures of all brain areas in these cohorts [52] and provides support for the use of visual scales as a sensitive marker of AD. However, the ability of the ≥1.5 MTA cut-off to classify healthy control subjects as not having the disease was not as high as previously reported in the same cohorts using automated methods [54]. In this regard, the age-based cut-off performed significantly better although at the cost of a lower sensitivity, suggesting that this is a more conservative approach.

Similar results were found in predicting the conversion from MCI to AD. In particular, a good

Table 5 *Classification performance of the ≥1.5 and age-dependent MTA cut-off scores in different age groups*

		≥1.5 cut-off		Age-dependent cut-off		≥2.0 cut-off	
	n	AD-like	CTR-like	AD-like	CTR-like	AD-like	CTR-like
50–74 years (%)							
AD	138	109 (79)	29 (21)	109 (79)	29 (21)	87 (63)	51 (37)
CTR	180	41 (22.8)	139 (77.2)	42 (23.3)	138 (76.7)	25 (13.9)	155 (86.1)
75–90 years (%)							
AD	184	163 (88.6)	21 (11.4)	116 (63)	68 (37)	150 (81.5)	34 (18.5)
CTR	165	69 (41.8)	96 (58.2)	16 (9.7)	149 (90.3)	42 (25.5)	123 (74.6)

AD, Alzheimer's disease; CTR, control.

sensitivity of 75.8% in supporting a diagnosis of AD in MCI subjects was achieved for individual cases as in the clinic by using a cut-off score of ≥1.5. However, the specificity based on this cut-off was lower, probably because many MCI subjects who remain stable after 1 year may, after a longer follow-up period, develop dementia. This might be the case for the MCI nonconverters of the AddNeuroMed cohort, who had low scores for immediate recall (ADAS1), similar to the MCI converters of both cohorts. By contrast, the age-corrected cut-off showed a significantly higher specificity for the prediction of MCI conversion, again at the cost of lower sensitivity. Hence, our results suggest that the selection of a specific cut-off plays a critical role in determining abnormality in MTA assessment. A lower threshold of ≥1.5 as used in previous studies [34, 35], although more sensitive in predicting the conversion of MCI to AD, might also misclassify many MCI subjects who do not convert to dementia as having an AD-like pattern.

To further investigate the differences between the two types of cut-offs, we compared the characteristics of the incorrectly and correctly classified subjects. We found that patients with AD, MCI nonconverters and control subjects who were incorrectly classified based on the ≥1.5 cut-off were significantly older and more often male compared to correctly classified individuals. By contrast, no significant differences in age or gender were found for the age-based classifications, suggesting that demographic variables have a greater impact on the ≥1.5 MTA cut-off. There is considerable evidence showing that normal ageing is associated with volume reductions in temporal areas and dilatation of the lateral ventricles [55–58] and that brain atrophy is more rapid in men, compared to women, with ageing [59–61]. Hence, increasing age and male gender might have contributed to the lower specificity of the ≥1.5 cut-off. Moreover, in addition to these variables, we observed worse MMSE and ADAS1 scores and a higher prevalence of the *ApoE* ε4 allele in the MCI nonconverters who were classified as having a more AD-like pattern based on the ≥1.5 cut-off. Previous studies have shown that the presence of the ε4 allele is one of the major genetic risk factors for AD [30, 62, 63], being associated with increased temporal brain atrophy both in AD and in MCI [29]. When applying the age-dependent cut-off, we observed that the only significant difference between correctly and incorrectly classified AD patients was a better memory performance in ADAS1 in those classified as having a more control-like pattern. It is possible that the age-dependent cut-off is sensitive to different stages of cognitive impairment in AD, discriminating patients who are more severely amnesic from those with better preserved cognitive function.

Due to the important role of *ApoE* status and disease onset in determining temporal atrophy in AD, we further assessed the separate and combined influence of these variables on the MTA cut-offs. Our findings showed a lower degree of temporal atrophy in patients with early-onset compared to those with late-onset AD based on the ≥1.5 cut-off. These results are consistent with the existence of a distinct disease phenotype in some patients with AD [64] who are typically diagnosed below 65 years of age and have a more posterior pattern of brain atrophy [26–28]. Using the age-dependent cut-off, there were no differences in MTA between patients with early- and late-onset disease, suggesting that this method can control the influence of disease onset on visual MTA. In addition, in contrast to the ≥1.5 cut-off, there was no significant difference in MTA between ε4 carriers and noncarriers using the age-dependent cut-off.

When patients with AD were further divided into four groups based on *ApoE* ε4 status and time of disease onset, this situation changed. Specifically, we observed significant differences between ε4 carriers and noncarriers using both MTA cut-offs, but only in patients with early-onset disease. On the other hand, we also found a higher degree of MTA in late-onset compared to early-onset patients, which was significant for the ≥1.5 cut-off with a trend towards significance for the age-dependent cut-off, but only in patients who did not carry the ε4 allele. These results suggest that carrying the ε4 allele and having a late onset of disease are associated with a higher degree of MTA using both cut-offs.

Finally, we also assessed the influence of age on MTA cut-offs by comparing the classification of patients with AD and control subjects at younger and older ages. We found that for both cut-offs, patients and controls were classified well between 50 and 74 years of age. However, from 75 to 90 years of age, the ≥1.5 cut-off showed low specificity, whilst the age-dependent cut-off showed low sensitivity, thus limiting the potential utility of both in the clinical setting. For this reason, a different cut-off of ≥2.0 was applied

because MTA is frequently observed in older age groups [20] and a higher degree of atrophy might be required to identify pathology. This cut-off correctly identified 81.5% of patients with AD and 74.5% of control subjects, exceeding the classification performances of the other two cut-offs.

This study has several strengths. First, we assessed MTA in the largest cohort to date based on 1147 MRI scans from the AddNeuroMed and ADNI studies. Secondly, we evaluated the performance of different MTA cut-offs, which might serve as a reference for future studies to determine a critical threshold of abnormality using the visual scale proposed by Schelten and colleagues.

Some limitations should also be recognized. First, visual assessment of MTA was only made using cross-sectional data because of the large number of subjects in the multicentre cohorts. Secondly, it would have been interesting to extend this work to include patients with other types of dementia such as frontotemporal dementia or Lewy bodies disease as MTA is not only present in AD. Such studies assessing the specificity of the visual scale proposed by Schelten *et al.* would be of significant clinical value. Thirdly, there was a lack of neuropathological confirmation in patients with AD in the present study; such confirmation would be helpful in establishing the utility of our findings in detecting both prodromal and dementia stages of AD. Fourthly, the lack of relevant neuropsychological data, such as delayed recall measures, assessed with the same version of ADAS-Cog for both cohorts is also a limitation of the current study. Finally, it would have been interesting to include visual measures of atrophy focused on posterior cortical areas, such as the recently developed posterior atrophy scale [65], which has already shown promising results in AD patients with early-onset disease. However, assessments of brain atrophy using the global cortical atrophy scale [66] and severity of white matter hyperintensity using the Fazekas scale [67], which are both easily implemented in a clinical setting, would have allowed us to perform a more detailed evaluation of the presence of atrophy in the current sample. Future studies assessing atrophy using all these visual scales in addition to other relevant biomarkers such as CSF levels of beta-amyloid and tau proteins will determine their value and how they can be combined to aid in the clinical assessment of patients with AD.

The goal of the current study was to evaluate the performance of both MTA cut-offs, in a large cohort of patients, with regard to their potential utility in a clinical setting. The decision to use one cut-off or the other lies with the clinician, and therefore here, we provide information on the sensitivity and specificity of each type of cut-off. Our findings show that an MTA cut-off of $\geq$1.5 is sensitive in detecting current and future conversion to AD, whereas an age-dependent cut-off is appropriate for more conservative classifications. In addition, our results strongly suggest the need to take into account demographic, genetic and clinical variables in routine clinical evaluations using the visual MTA scale. Older healthy control subjects can experience substantial MTA, and for this reason, we suggest using a higher threshold of $\geq$2.0 to determine abnormality after 75 years of age. On the other hand, the diagnosis of AD in the absence of *ApoE* ε4 or in patients with early-onset disease should be based on clinical information extending beyond MTA ratings, as MTA may not be clearly present in these cases. Hence, we recommend using visual MTA assessment if a patient has a late onset of disease and carries at least one *ApoE* ε4 allele. If the patient is younger than 75 years of age, a cut-off score of $\geq$1.5 is more appropriate, whereas a higher threshold of $\geq$2.0 is more accurate from 75 years onwards. In the case of significant doubt about the clinical diagnosis and preference for a more conservative approach, an age-based cut-off offers a higher specificity. Other variables such as measures of immediate recall are also worth assessing as patients with better memory skills presented a lower degree of atrophy. Finally, gender also plays a relevant role, as men may have a higher degree of MTA than women; clinicians should be aware of this when using the visual MTA scale as part of the AD diagnostic workup.

Conflict of interest statement

No conflict of interest to declare.

Acknowledgements

This study was supported by InnoMed (Innovative Medicines in Europe), an Integrated Project funded by the European Union of the Sixth Framework programme priority FP6-2004-LIFESCIHEALTH-5, Life Sciences, Genomics and Biotechnology for Health. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) [National Institutes of Health

(NIH) 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: Alzheimer's Association, Alzheimer's Drug Discovery Foundation, BioClinica Inc., Biogen Idec Inc., Bristol-Myers Squibb Co., Eisai, Inc., Elan Pharmaceuticals, Inc., Eli Lilly and Co., F. Hoffmann-La Roche Ltd and its affiliated company Genentech Inc., GE Healthcare, Innogenetics NV, IXICO Ltd, Janssen Alzheimer Immunotherapy Research & Development LLC, Johnson & Johnson Pharmaceutical Research & Development LLC, Medpace Inc., Merck & Co. Inc. Meso Scale Diagnostics LLC, NeuroRx Research, Novartis Pharmaceuticals Corp., Pfizer Inc., Piramal Imaging, Servier, Synarc Inc. and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's 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 and K01 AG030514.

The authors thank Swedish Brain Power, the Strategic Research Programme in Neuroscience at Karolinska Institutet (StratNeuro), the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, Health Research Council of Academy of Finland, University of Eastern Finland UEFBRAIN and EVO funding from Kuopio University Hospital. AS and SL were supported by funds from National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health and the NIHR Biomedical Research Unit for Dementia at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, Kings College London. JBP was funded by a Marie Curie fellowship for postdoctoral researchers (FP7-PEOPLE-2012-IEF).

References

1 Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med* 2004; **256:** 183–94.
2 Ritchie K, Artero S, Touchon J. Classification criteria for mild cognitive impairment: a population-based validation study. *Neurology* 2001; **56:** 37–42.
3 Petersen RC, Doody R, Kurz A *et al.* Current concepts in mild cognitive impairment. *Arch Neurol* 2001; **58:** 1985–92.
4 DeCarli C, Frisoni GB, Clark CM *et al.* Qualitative estimates of medial temporal atrophy as a predictor of progression from mild cognitive impairment to dementia. *Arch Neurol* 2007; **64:** 108–15.
5 Duara R, Loewenstein DA, Potter E *et al.* Medial temporal lobe atrophy on MRI scans and the diagnosis of Alzheimer disease. *Neurology* 2008; **71:** 1986–92.
6 Dickerson BC, Goncharova I, Sullivan MP *et al.* MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. *Neurobiol Aging* 2001; **22:** 747–54.
7 Du AT, Schuff N, Amend D *et al.* Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2001; **71:** 441–7.
8 Chetelat G, Landeau B, Eustache F *et al.* Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: a longitudinal MRI study. *Neuroimage* 2005; **27:** 934–46.
9 Wahlund LO, Julin P, Johansson SE, Scheltens P. Visual rating and volumetry of the medial temporal lobe on magnetic resonance imaging in dementia: a comparative study. *J Neurol Neurosurg Psychiatry* 2000; **69:** 630–5.
10 Jack CR, Petersen RC, Xu Y *et al.* Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. *Neurology* 2000; **55:** 484–9.
11 Killiany RJ, Hyman BT, Gomez-Isla T *et al.* MRI measures of entorhinal cortex vs hippocampus in preclinical AD. *Neurology* 2002; **58:** 1188–96.
12 Devanand DP, Pradhaban G, Liu X *et al.* Hippocampal and entorhinal atrophy in mild cognitive impairment: prediction of Alzheimer disease. *Neurology* 2007; **68:** 828–36.
13 Risacher SL, Saykin AJ, West JD, Shen L, Firpi HA, McDonald BC. Alzheimer's Disease Neuroimaging Initiative. Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. *Curr Alzheimer Res* 2009; **6:** 347–61.
14 Scheltens P, Leys D, Barkhof F *et al.* Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *J Neurol Neurosurg Psychiatry* 1992; **55:** 967–72.
15 Visser PJ, Verhey FR, Hofman PA, Scheltens P, Jolles J. Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. *J Neurol Neurosurg Psychiatry* 2002; **72:** 491–7.
16 Korf ES, Wahlund LO, Visser PJ, Scheltens P. Medial temporal lobe atrophy on MRI predicts dementia in patients with mild cognitive impairment. *Neurology* 2004; **63:** 94–100.
17 Bouwman FH, Schoonenboom SN, van der Flier WM *et al.* CSF biomarkers and medial temporal lobe atrophy predict dementia in mild cognitive impairment. *Neurobiol Aging* 2007; **28:** 1070–4.
18 Burton EJ, Barber R, Mukaetova-Ladinska EB *et al.* Medial temporal lobe atrophy on MRI differentiates Alzheimer's disease from dementia with Lewy bodies and vascular cognitive impairment: a prospective study with pathological verification of diagnosis. *Brain* 2009; **132:** 195–203.

19 van de Pol LA, Korf ES, van der Flier WM *et al.* Magnetic resonance imaging predictors of cognition in mild cognitive impairment. *Arch Neurol* 2007; **64:** 1023–8.
20 Launer LJ, Scheltens P, Lindeboom J, Barkhof F, Weinstein HC, Jonker C. Medial temporal lobe atrophy in an open population of very old persons: cognitive, brain atrophy, and sociomedical correlates. *Neurology* 1995; **45:** 747–52.
21 Galton CJ, Patterson K, Graham K *et al.* Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology* 2001; **57:** 216–25.
22 Barber R, Ballard C, McKeith IG, Gholkar A, O'Brien JT. MRI volumetric study of dementia with Lewy bodies: a comparison with AD and vascular dementia. *Neurology* 2000; **54:** 1304–9.
23 Tam CWC, Burton EJ, McKeith IG, Burn DJ, O'Brien JT. Temporal lobe atrophy on MRI in Parkinson's disease with dementia: a comparison with Alzheimer's disease and dementia with Lewy bodies. *Neurology* 2005; **64:** 861–5.
24 Galton CJ, Patterson K, Xuereb JH, Hodges JR. Atypical and typical presentations of Alzheimer's disease: a clinical, neuropsychological, neuroimaging and pathological study of 13 cases. *Brain* 2000; **3:** 484–98.
25 Kemp PM, Holmes C, Hoffmann SM *et al.* Alzheimer's disease: differences in technetium-99 m HMPAO SPECT scan findings between early onset and late onset dementia. *J Neurol Neurosurg Psychiatry* 2003; **74:** 715–9.
26 Jones BF, Barnes J, Uylings HB *et al.* Differential regional atrophy of the cingulate gyrus in Alzheimer disease: a volumetric MRI study. *Cereb Cortex* 2006; **16:** 1701–8.
27 Karas G, Scheltens P, Rombouts S *et al.* Precuneus atrophy in early-onset Alzheimer's disease: a morphometric structural MRI study. *Neuroradiology* 2007; **49:** 967–76.
28 Whitwell JL, Dickson DW, Murray ME *et al.* Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. *Lancet Neurol* 2012; **11:** 868–77.
29 Cherbuin N, Leach LS, Christensen H, Anstey KJ. Neuroimaging and *ApoE* genotype: a systematic qualitative review. *Dement Geriatr Cogn Disord* 2007; **24:** 348–62.
30 Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat Rev Neurol* 2013; **9:** 106–18.
31 Jack CR, Petersen RC, Xu YC *et al.* Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology* 1997; **49:** 786–94.
32 van de Pol LA, Hensel A, van der Flier WM *et al.* Hippocampal atrophy on MRI in frontotemporal lobar degeneration and Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2006; **77:** 439–42.
33 Barkhof F, Polvikoski TM, van Straaten EC *et al.* The significance of medial temporal lobe atrophy: a postmortem MRI study in the very old. *Neurology* 2007; **69:** 1521–7.
34 Schoonenboom NS, van der Flier WM, Blankenstein MA *et al.* CSF and MRI markers independently contribute to the diagnosis of Alzheimer's disease. *Neurobiol Aging* 2008; **29:** 669–75.
35 van Rossum IA, Visser PJ, Knol DL *et al.* Injury markers but not amyloid markers are associated with rapid progression from mild cognitive impairment to dementia in Alzheimer's disease. *J Alzheimers Dis* 2012; **29:** 319–27.
36 Westman E, Simmons A, Muehlboeck JS *et al.* AddNeuroMed and ADNI: similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America. *Neuroimage* 2011; **58:** 818–28.
37 Lehman M, Koedam EL, Barnes J *et al.* Visual ratings of atrophy in MCI: prediction of conversión and relationship with CSF biomarkers. *Neurobiol Aging* 2013; **34:** 73–82.
38 Lovestone S, Francis P, Strandgaard K. Biomarkers for disease modification trials–the innovative medicines initiative and AddNeuroMed. *J Nutr Health Aging* 2007; **11:** 359–61.
39 Lovestone S, Francis P, Kloszewska I *et al.* AddNeuroMed – the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. *Ann N Y Acad Sci* 2009; **1180:** 36–46.
40 McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS Alzheimer's disease RDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology* 1984; **34:** 939–44.
41 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; **12:** 189–98.
42 Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984; **141:** 1356–64.
43 Shaw LM, Vanderstichele H, Knapik-Czajka M *et al.* Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol* 2009; **65:** 403–13.
44 Hixson JE, Vernier DT. Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with HhaI. *J Lipid Res* 1990; **31:** 545–8.
45 Jack CR, Bernstein MA, Fox NC *et al.* The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging* 2008; **27:** 685–91.
46 Simmons A, Westman E, Muehlboeck S *et al.* MRI measures of Alzheimer's disease and the AddNeuroMed study. *Ann N Y Acad Sci* 2009; **1180:** 47–55.
47 Simmons A, Westman E, Muehlboeck S *et al.* The AddNeuroMed framework for multi-centre MRI assessment of Alzheimer's disease: experience from the first 24 months. *Int J Geriatr Psychiatry* 2011; **26:** 75–82.
48 Cavallin L, Bronge L, Zhang Y *et al.* Comparison between visual assessment of MTA and hippocampal volume in an elderly, non-demented population. *Acta Radiol* 2012; **53:** 573–9.
49 Hawass NE. Comparing the sensitivities and specificities of two diagnostic procedures performed on the same group of patients. *Br J Radiol* 1997; **70:** 360–6.
50 Jonker C, Geerlings MI, Schmand B. Are memory complaints predictive for dementia? A review of clinical and population-based studies. *Int J Geriatr Psychiatry* 2000; **15:** 983–91.
51 Braak H, Braak E. Alzheimer's disease affects limbic nuclei of the thalamus. *Acta Neuropathol* 1991; **81:** 261–8.
52 Dubois B, Feldman HH, Jacova C *et al.* Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 2007; **6:** 734–46.
53 Albert MS, DeKosky ST, Dickson D *et al.* The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011; **7:** 270–9.
54 Spulber G, Simmons A, Muehlboeck JS *et al.* An MRI-based index to measure the severity of Alzheimer's disease-like

structural pattern in subjects with mild cognitive impairment. *J Intern Med* 2013; **273:** 396–409.
55 Raz N, Gunning-Dixon F, Head D, Rodrigue KM, Williamson A, Acker JD. Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. *Neurobiol Aging* 2004; **25:** 377–96.
56 Salat DH, Buckner RL, Snyder AZ *et al.* Thinning of the cerebral cortex in aging. *Cereb Cortex* 2004; **14:** 721–30.
57 Walhovd KB, Fjell AM, Reinvang I *et al.* Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiol Aging* 2005; **26:** 1261–70.
58 Fjell AM, Walhovd KB, Fennema-Notestine C *et al.* One-year brain atrophy evident in healthy aging. *J Neurosci* 2009; **29:** 15223–31.
59 Xu J, Kobayashi S, Yamaguchi S, Iijima K, Okada K, Yamashita K. Gender effects on age-related changes in brain structure. *AJNR Am J Neuroradiol* 2000; **21:** 112–8.
60 Sowell ER, Peterson BS, Kan E *et al.* Sex differences in cortical thickness mapped in 176 healthy individuals between 7 and 87 years of age. *Cereb Cortex* 2007; **17:** 1550–60.
61 Pfefferbaum A, Rohlfing T, Rosenbloom MJ, Chu W, Colrain IM, Sullivan EV. Variation in longitudinal trajectories of regional brain volumes of healthy men and women (ages 10 to 85 years) measured with atlas-based parcellation of MRI. *Neuroimage* 2013; **65:** 176–93.
62 Raber J, Huang Y, Ashford JW. *ApoE* genotype accounts for the vast majority of AD risk and AD pathology. *Neurobiol Aging* 2004; **25:** 641–50.
63 Rogalski EJ, Rademaker A, Harrison TM *et al. ApoE* E4 is a susceptibility factor in amnestic but not aphasic dementias. *Alzheimer Dis Assoc Disord* 2011; **25:** 159–63.
64 van der Flier WM, Pijnenburg YA, Fox NC, Scheltens P. Early-onset versus late-onset Alzheimer's disease: the case of the missing *ApoE* varepsilon4 allele. *Lancet Neurol* 2011; **10:** 280–8.
65 Lehmann M, Koedam EL, Barnes J *et al.* Posterior cerebral atrophy in the absence of medial temporal lobe atrophy in pathologically-confirmed Alzheimer's disease. *Neurobiol Aging* 2012; **33:** 627.e1–e12.
66 Pasquier F, Leys D, Weerts JG, Mounier-Vehier F, Barkhof F, Scheltens P. Inter- and intraobserver reproducibility of cerebral atrophy assessment on MRI scans with hemispheric infarcts. *Eur Neurol* 1996; **36:** 268–72.
67 Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *Am J Roentgenol* 1987; **149:** 351–6.

Correspondence: Eric Westman, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Novum 5th floor, SE-141 86 Stockholm, Sweden.
(fax: +46 858585470; e-mail: eric.westman@ki.se).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Comparisons between the MTA models.

