Neurobiology of Aging 34 (2013) 2003–2013

Contents lists available at SciVerse ScienceDirect

Neurobiology of Aging

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

Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment

Lies Clerx [a,*], Ineke A. van Rossum [b], Leah Burns [c], Dirk L. Knol [d], Philip Scheltens [b], Frans Verhey [a], Pauline Aalten [a], Pablo Lapuerta [e], Laura van de Pol [b], Ronald van Schijndel [b], Remko de Jong [b], Frederik Barkhof [b], Robin Wolz [f], Daniel Rueckert [f], Martina Bocchetta [g], Magdalini Tsolaki [h], Flavio Nobili [i], Lars-Olaf Wahlund [j], Lennart Minthon [k], Lutz Frölich [l], Harald Hampel [m], Hilkka Soininen [n], Pieter Jelle Visser [a,b]

[a] Department of Psychiatry and Neuropsychology, Maastricht University, School for Mental Health and Neuroscience, Alzheimer Center Limburg, Maastricht, the Netherlands
[b] Department of Neurology, Alzheimer Center, VU University Medical Center, Amsterdam, the Netherlands
[c] Global Health Economics & Outcomes Research, Bristol-Myers Squibb Company, Princeton, NJ, USA
[d] Department of Clinical Epidemiology and Statistics, VU University Medical Center, Amsterdam, the Netherlands
[e] Lexicon Pharmaceuticals, The Woodlands, TX, USA
[f] Department of Computing, Imperial College London, London, UK
[g] IRCCS San Giovanni, Laboratory of Epidemiology and Neuroimaging, Brescia, Italy
[h] Aristotle University of Thessaloniki, Memory and Dementia Centre, 3rd Department of Neurology, G Papanicolaore General Hospital, Thessaloniki, Greece
[i] Clinical Neurophysiology Service, Department of Neurosciences, Ophthalmology and Genetics, University of Genoa, Genoa, Italy
[j] NVS Department, Karolinska Institutet, Karolinska University Hospital, Huddinge, Sweden
[k] Clinical Memory Research Unit, Department of Clinical Sciences, Malmö Lund University, Malmö, Sweden
[l] Department of Geriatric Psychiatry, Zentralinstitut für Seelische Gesundheit, University of Heidelberg, Mannheim, Germany
[m] Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, Goethe-University, Frankfurt, Germany
[n] Department of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland

ARTICLE INFO

Article history:
Received 24 January 2013
Accepted 9 February 2013
Available online 27 March 2013

Keywords:
Alzheimer's disease
Mild cognitive impairment
Dementia
Magnetic resonance imaging
Medial temporal lobe atrophy
Cerebrospinal fluid (CSF)
Aβ1–42
Tau

ABSTRACT

Our aim was to compare the predictive accuracy of 4 different medial temporal lobe measurements for Alzheimer's disease (AD) in subjects with mild cognitive impairment (MCI). Manual hippocampal measurement, automated atlas-based hippocampal measurement, a visual rating scale (MTA-score), and lateral ventricle measurement were compared. Predictive accuracy for AD 2 years after baseline was assessed by receiver operating characteristics analyses with area under the curve as outcome. Annual cognitive decline was assessed by slope analyses up to 5 years after baseline. Correlations with biomarkers in cerebrospinal fluid (CSF) were investigated. Subjects with MCI were selected from the Development of Screening Guidelines and Clinical Criteria for Predementia AD (DESCRIPA) multicenter study ($n = 156$) and the single-center VU medical center ($n = 172$). At follow-up, area under the curve was highest for automated atlas-based hippocampal measurement (0.71) and manual hippocampal measurement (0.71), and lower for MTA-score (0.65) and lateral ventricle (0.60). Slope analysis yielded similar results. Hippocampal measurements correlated with CSF total tau and phosphorylated tau, not with beta-amyloid 1–42. MTA-score and lateral ventricle volume correlated with CSF beta-amyloid 1–42. We can conclude that volumetric hippocampal measurements are the best predictors of AD conversion in subjects with MCI.

1. Introduction

Alzheimer's disease (AD) is the leading cause of dementia in the elderly, affecting more than 27 million people worldwide. Early detection of AD might prevent irreversible damage by enabling preventative treatment (Masters and Beyreuther, 2006; Vellas et al., 2007). A primary focus of research in AD is identifying which biomarkers are clinically useful for the early diagnosis of AD.

Medial temporal lobe (MTL) atrophy as assessed using structural magnetic resonance imaging (MRI) has proven to be an effective clinical aid in the early diagnosis of AD (Visser et al., 2002a), and this method predicts AD in subjects with mild cognitive impairment (MCI) (DeCarli et al., 2007; Rusinek et al., 2004; Schoonenboom

* Corresponding author at: Department of Psychiatry and Neuropsychology, Maastricht University Medical Center, PO Box 616, 6200 MD Maastricht, the Netherlands. Tel.: +31 43 3881841; fax: +31 43 3884092.
E-mail address: L.Clerx@maastrichtuniversity.nl (L. Clerx).

0197-4580
http://dx.doi.org/10.1016/j.neurobiolaging.2013.02.002

et al., 2008; Visser et al., 1999, 2002a). There are several ways to determine the degree of MTL atrophy in the brain including manual delineation, (semi-) automated techniques to measure hippocampal volume, qualitative ratings of MTL atrophy (MTA-score), and assessment of lateral ventricular volume (Aljabar et al., 2009; Apostolova et al., 2006; Chou et al., 2010; Echávarri et al., 2011; Jack et al., 2004; McHugh et al., 2007; Nestor et al., 2008; Scheltens et al., 1992; Varela-Nallar et al., 2010; Wolz et al., 2010a, 2010b). Each of these methods has its strengths and limitations.

Manual volumetry is considered the gold standard (Barnes et al., 2009; Boccardi et al., 2011; van de Pol et al., 2007b) but is time-consuming, which limits routine clinical or large-scale research use. Automated measurements are quick and widely applicable, but might be susceptible to scanner and scan protocol variability. Volumetric measurements of the lateral ventricle require a minimum of rater time with robust automatic segmentations but show a lot of variability and asymmetry between subjects (Nestor et al., 2008). Qualitative ratings are quick to perform but sensitive to interrater variability and show lower accuracy rates compared with volumetric analysis (DeCarli et al., 1990; Galton et al., 2001). Visual rating scales are furthermore insensitive to detect atrophy progression over time (Ridha et al., 2007).

A number of studies have examined differences between various techniques to measure atrophy of the MTL, mostly comparing manual with automated hippocampal volumetry (Lehmann et al., 2010; Sanchez-Benavides et al., 2010b; Shen et al., 2010) or volumetric hippocampal measurements to a visual rating scale (Ridha et al., 2007; Scheltens et al., 1992; Urs et al., 2009; Wahlund et al., 2000; Westman et al., 2011). These studies typically evaluate the diagnostic accuracy of different MRI techniques by comparing AD patients with healthy control subjects. Most studies found that manual hippocampal measurement and automated hippocampal segmentation results were similar (Hsu et al., 2002; Lehmann et al., 2010). However, the performance of automated techniques might be less precise when applied in AD patients suffering from moderate to severe brain atrophy and/or white matter hyperintensities which might lead to false allocations of gray matter, white matter, or cerebrospinal fluid (CSF) (Carmichael et al., 2005; Levy-Cooperman et al., 2008; Sanchez-Benavides et al., 2010a). One study reported that visual rating of MTL atrophy is a quick and clinically useful technique for differentiating AD from control subjects and is quicker and more accurate than volumetry (Wahlund et al., 2000).

To our knowledge, no study has compared the diagnostic performance of manual and atlas-based hippocampal segmentation, lateral ventricle volume, and a qualitative rating. Moreover, no comparative studies have been performed on the predictive accuracy of these different methods to predict AD in subjects with MCI, their relation with CSF biomarkers of AD, and the effect of multicenter settings on diagnostic performance.

The aim of the present longitudinal study was to compare the predictive accuracy of 4 different MTL measurements for the progression to AD-type dementia in patients with MCI over a 2-year follow-up period. Atrophy of the MTL was assessed using manual measurement of the hippocampus, automatically measured hippocampal volume based on atlas registration (learning embeddings for atlas propagation; LEAP), volumetric measurement of the expansion of the lateral ventricle, and a largely used qualitative rating scale. Because subjects might convert at a later point in time, slope analyses were additionally performed with annual cognitive decline up to 5 years as an outcome measure. The correlation of MTL measures with AD biomarkers in CSF was also investigated and the predictive accuracy was tested in a multicenter study with different scan protocols and in a single-center study.

2. Materials and methods

2.1. Subjects

We selected participants with MCI from the Development of Screening Guidelines and Clinical Criteria for Predementia AD (DESCRIPA) study and the Alzheimer Center of the VU University Medical center (VUmc) in Amsterdam. DESCRIPA is a multicenter prospective cohort study from the European Alzheimer's Disease Consortium aimed at developing clinical criteria and screening guidelines for predementia AD (Visser et al., 2008). For this study participants were selected from 9 of the 20 participating centers where MRI scanning was performed as part of clinical practice or as research protocol. The VUmc cohort in the present study included the VUmc subjects enrolled in the DESCRIPA study and an additional sample of subjects that were seen outside the DESCRIPA inclusion period (Supplementary Appendix 1).

Inclusion criteria for both cohorts were: age 54 years or older, diagnosis of MCI, and availability of results for each MRI measure and outcome at follow-up. Exclusion criteria were diagnosis of dementia at baseline or any somatic, psychiatric, or neurological disorder (e.g., epilepsy) that might have caused the cognitive impairment (Visser et al., 2008). At baseline, scans were available for 456 subjects. Visually rated MTL atrophy was available for all subjects. Of these, 54 had no follow-up data and were excluded. Of the remaining 402 subjects, scans were not available in digital format for 21 subjects. Of the remaining 381 scans, manual segmentation of the hippocampus could be performed on 341 scans (reasons missing: technical problem in volumetric measurement [$n = 5$], technical problem in automated intracranial volume estimation [$n = 25$], and logistical [$n = 10$]), LEAP-based volumetry on 357 scans (reasons missing: technical problem in volumetric measurement [$n = 11$] and logistical [$n = 13$]), and lateral ventricle volumetry on 335 scans (reasons missing: technical problem in volumetric measurement [$n = 37$], and logistical [$n = 9$]). Data for all 4 medial temporal lobe measurements were available for 328 subjects; 156 from DESCRIPA and 172 from VUmc. There were no differences between included and excluded subjects with respect to age, sex, educational level, and cognitive test results.

2.2. Clinical and cognitive assessment

All participants underwent a standard diagnostic workup, including clinical history, medical and neurological examination, clinical chemistry, functional evaluation using the Clinical Dementia Rating scale (Morris, 1993), the Mini-Mental State Examination (MMSE), and rating scales for depression and neuropsychiatric symptoms. A neuropsychological battery was performed to evaluate performance in several cognitive domains. In each center a primary test for verbal memory, language, attention, executive functioning, and visuoconstruction was chosen that was identical or similar to tests used in other centers (Visser et al., 2008). Raw scores on neuropsychological tests were corrected for age, education, and sex, in accordance with locally collected or published normative data and expressed as z-scores, which were used for further analysis. Baseline diagnosis of MCI was made according to the criteria of Petersen and colleagues (Petersen, 2004; Petersen et al., 1999). Subjects with a z-score $<$-1.5 SD on any of the following tests: the learning measure or delayed recall of a word list learning test or equivalent memory test, the Trail Making Test part A, Trail Making Test part B, verbal fluency, Rey figure copy test or an equivalent test were defined as having MCI (reference Vos). (Vos et al., 2012). We calculated a composite score as the average z-score of the 6 tests if scores were available for at least 3 tests (Visser et al., 2009).

Follow-up was conducted annually for up to 5 years. The primary outcome measure was conversion to AD-type dementia

after 2 years. AD diagnosis was made according to the *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition (American Psychiatric Association, 1994) and National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association criteria (McKhann et al., 1984). Secondary outcome measures were annual cognitive decline on the MMSE and the cognitive composite score.

2.3. MRI acquisition and image analysis

2.3.1. Scan protocol

At each site, patients were scanned according to the routine MRI protocol. Scanners and protocols at different sites varied but all scanning was performed at 1.0 or 1.5 Tesla (Supplementary Appendix 1). All scans included a 3-dimensional T1-weighted gradient echo sequence and a fast fluid-attenuated inversion recovery sequence.

2.3.2. MRI measurements

MTL atrophy was assessed using manual measurement of the hippocampus, automatically measured hippocampal volume using multi-atlas segmentation (LEAP), volumetric measurement of the expansion of the lateral ventricle, and a qualitative rating. All volumetric measurements were corrected for intracranial volume (ICV). The LEAP measurement was ICV-corrected by means of a scaling factor from Montreal Neurological Institute (MNI), the manual hippocampal measurement and the lateral ventricle measurement were ICV-corrected by means of a scaling factor derived from FSL software (FMRIB, Oxford, UK). The total rating time of each method can be found in Supplementary Table 1. Rater time needed for analysis was lowest for LEAP hippocampal measurement (4 minutes) followed by the qualitative rating (5 minutes). Rater time was 30 minutes for lateral ventricle measurement and 150 minutes for manual hippocampal volumetry.

For the manual segmentation of the hippocampus, the baseline 3-dimensional T1-weighted volume sequence was reformatted in 2-mm slices (in-plane resolution: 1×1 mm) and oriented perpendicular to the long axis of the left hippocampus (van de Pol et al., 2007a). Regions of interest (ROIs) of the hippocampus were constructed by manual delineation of hippocampal borders on both sides on the reformatted slices, using the software package developed in-house, Show_Images 3.7.0 (VU University Medical Center, Amsterdam, the Netherlands). Delineation of the hippocampus was performed using previously described criteria (Jack, 1994; van de Pol et al., 2007a, 2009) by 3 trained technicians (coefficients of variation: interrater <8%, intrarater <5%) blinded to diagnosis. ROIs included the dentate gyrus, cornu ammonis, subiculum, fimbria, and alveus. Baseline hippocampal volume was calculated by multiplying the total area of all ROIs of each hippocampus by slice thickness. Baseline hippocampal volumes were adjusted for ICV, using the scaling factor derived from SIENAX (part of FSL; FMRIB) (Sluimer et al., 2008).

Automated hippocampal volumetry was performed using the LEAP method (Wolz et al., 2010a; Supplementary Fig. 1). In this method, multi-atlas registration is applied to a cohort of brain images after representing the whole population together with an initial set of atlases. The initial set is propagated to a number of unlabeled images in their local neighborhood which are used to label them. Images labeled in this way become atlases themselves and are, consequently, further propagated throughout the whole data set. In this way, each image is labeled using a number of atlases in its close vicinity, which has been shown to perform more robustly on diverse data sets than other multi-atlas registration techniques (Wolz et al., 2010a). A brief visual inspection of the segmented hippocampi was performed to identify clear failures of the automated method. Except for technical failures listed below (see 4.3. Technical considerations), no subjects were excluded after this inspection and no manual correction was performed.

Measurement of the lateral ventricle was executed with an extension of SIENAX (Smith et al., 2001, 2002), part of FSL (FMRIB) (Smith et al., 2004). SIENAX starts by extracting brain and skull images from the single whole-head input data (Smith, 2002). The brain image is then affine-registered to MNI152 space (Jenkinson and Smith, 2001; Jenkinson et al., 2002) using the skull image to determine the registration scaling. This is primarily to obtain the volumetric scaling factor, to be used as a normalization for head size. Next, tissue-type segmentation is carried out (Zhang et al., 2001) to calculate total volume of brain tissue (including separate estimates of volumes of gray matter, white matter, peripheral gray matter, and ventricular CSF). After the tissue segmentation, a registered mask is used to exclude the CSF on the outer side of the brain. The resulting ventricular structure is manually edited when the segmentation or the mask did not take the whole ventricle into account or contained CSF pixels not belonging to the ventricles (CSF found outside the ventricles).

The visual rating of MTL atrophy was performed using a qualitative scale (Scheltens et al., 1992). Rating was performed on coronal T1-weighted images using a 5-point visual scale (MTA-scores), ranging from 0 (no atrophy) to 4 (severe atrophy) based on the height of the hippocampal formation and the surrounding CSF spaces. In the analysis, the sum of both sides (left and right) was used. All visual rating was performed at VUmc by a number of trained raters. Scans from VUmc were rated by a group of 3 raters supervised by a neuroradiologist (intrarater weighted Cohen $\kappa >$ 0.80; interrater weighted Cohen $\kappa > 0.80$) (Henneman et al., 2009). The DESCRIPA scans were rated by a single rater from VUmc (intrarater weighted Cohen $\kappa = 0.68$) (van de Pol et al., 2009).

2.4. CSF analyses

CSF was collected by lumbar puncture and levels of beta amyloid (Aβ)1–42, total tau (t-tau), and phosphorylated tau (p-tau) in CSF were measured using commercially available sandwich enzyme-linked immunosorbent assays (Innotest β-amyloid 1-42; Innotest hTAU-Ag; Innogenetics, Ghent, Belgium), specially constructed to measure Aβ1–42 and t-tau (Andreasen et al., 1999; Blennow et al., 1995) by experienced technicians at the lab in Gothenburg for the DESCRIPA cohort and in Amsterdam for the VUmc cohort. CSF was available for 147 subjects. We corrected for interlaboratory enzyme-linked immunosorbent assay differences by analyzing 33 samples at both labs and we adjusted VUmc values to those of DESCRIPA using the following equating formula: Gothenborg = average Gothenborg + (SD Gothenborg/SD VUmc) × (VUmc − average VUmc) (Kolen and Brennan, 1995). Abnormal values for CSF measures were a concentration ≤550 pg/mL for Aβ1–42, >52 pg/mL for p-tau, and >375 pg/mL for t-tau (Mulder et al., 2010).

2.5. APOE *genotype*

APOE genotype was determined by polymerase chain reaction of genomic DNA extracted from ethylenediaminetetraacetic acid (EDTA)–anticoagulated blood in 262 subjects. Subjects were classified as *APOE*-ε4 carriers or noncarriers.

2.6. Statistical analysis

Statistical analysis was performed with IBM SPSS version 19 (Chicago, IL, USA) and the statistical software package R (R Foundation, Vienna, Austria). Correlations between the different methods for MTL assessment and between MTL and other biomarkes (CSF, *APOE*-e4 allele) were analyzed by the Pearson

correlation coefficient for continuous data, and the Spearman rank test for correlations including the MTA-score. Intraclass correlation coefficient (ICC) and paired *t* test were performed to investigate the agreement between MTL measurements.

The main outcome measure was the area under the curve (AUC) for AD-type dementia after 2 years of follow-up, calculated using a time-dependent receiver operating characteristics curve in R (Heagerty et al., 2000). Differences in AUC between methods were tested as described elsewhere (Hanley and McNeil, 1983). We calculated the sensitivity, specificity, positive predictive value, negative predictive value, odds ratio (OR), and hazard ratio (HR) for AD-type dementia at the 2-year follow-up using data-driven cut points based on a time-dependent receiver operating characteristics analysis. First we calculated the cut point that maximized the Youden Index (sensitivity + specificity − 1) for predicting AD-type dementia after 2 years. Second, we selected cut points that predicted AD-type dementia with a sensitivity of 85%. Cut points were calculated in the whole sample and in each cohort separately. Spline analyses were performed to determine AD-free survival after 2 years as a function of each MRI measurement in the total cohort. Slope analyses with mixed models were performed to investigate whether MTL atrophy was associated with change on the MMSE and a cognitive composite score at follow-up. The analyses included the baseline score and available follow-up scores up to 5 years after baseline and were corrected for age and education. An unstructured covariance structure with center as a random effect was used because this model provided the best −2 log likelihood compared with models with simpler covariance structures (Visser et al., 2009).

3. Results

3.1. Subject characteristics

From the total group of MCI patients included in this study 37% were diagnosed as nonamnestic MCI (73% single-domain and 27% multi-domain) and 63% amnestic MCI (44% single-domain and 56% multi-domain). Baseline and follow-up characteristics are shown in Table 1. Both cohorts were comparable for age and *APOE*-e4 status. The DESCRIPA cohort included more female participants and education was lower compared with the VUmc cohort. Scores on the MMSE and delayed recall task were lower in the VUmc cohort. At follow-up, 91 subjects were diagnosed with probable Alzheimer-type dementia (28%). Conversion rate was higher in the VUmc cohort (35.5%) than in the DESCRIPA cohort (19.2%), $p < 0.001$. Twelve subjects converted to a different type of dementia at follow-up: 4 subjects converted to frontotemporal dementia, 6 to dementia with Lewy bodies, 1 to vascular dementia, and 1 to another form of dementia. These subjects were included in the no-AD group. The follow-up length was slightly longer in the VUmc cohort. Characteristics of subjects with and without AD-type dementia at follow-up in each cohort are shown in Supplementary Table 2.

3.2. Correlations between MTL measures

Manual and automated LEAP hippocampal volumes correlated highly (Pearson $r = 0.71$; $p < 0.001$) with intermediate agreement (ICC, 0.38; $p < 0.001$, Bland–Altman plot; Supplementary Fig. 2). The 2 techniques were significantly different using the paired *t* test ($t = 42.27$; $p < 0.01$) with the LEAP hippocampal measurement showing lower volumes. Using the cut point based on the Youden index, the κ and the overlap between manual and LEAP hippocampal volume were 0.60 and 0.80 respectively.

The correlation between the MTA-score and the manual and LEAP hippocampal volume was −0.36 and −0.27 respectively ($p < 0.01$). Using the cut point based on the Youden index, the κ was 0.29 and the overlap between scores 0.69 for manual versus MTA-score and 0.37 (κ) and 0.65 (overlap) for LEAP versus MTA-score. The lateral ventricle measurement strongly correlated with MTA-score ($r = 0.60$; $p < 0.001$) and showed weak correlations with manual ($r = -0.20$; $p < 0.01$) and LEAP hippocampal volume ($r = -0.20$; $p < 0.01$).

Table 1
Subject characteristics

	DESCRIPA cohort	VUmc cohort	Combined cohort
n	156	172	328
Age, y	70.3 (7.9)	70.9 (7.3)	70.6 (7.6)
Female (%)	59.6	44.2[a]	51.5
Education, y	8.5 (3.9)	11.2 (3.3)[b]	10 (3.8)
MMSE score	27.2 (2.3)	26.6 (2.6)[c]	27.0 (2.5)
Z-score word list (delayed recall)	−1.23 (1.20)	−1.58 (1.10)[c]	−1.38 (1.20)
ApoE-e4 carrier (%)	45	56	51
Homo/heterozygous *ApoE*-e4 carriers (%)	38/7	38/18	38/13
Manual hippocampus (mm^3)	7874 (1294)	7462 (1097)[a]	7657 (1211)
LEAP hippocampus (mm^3)	5897 (798)	5446 (632)[b]	5661 (749)
MTA-score	2.8 (1.7)	2.3 (1.7)[a]	2.5 (1.7)
Lateral ventricle (mm^3)	53,409 (26,105)	57,374 (28,042)	55,488 (27,170)
CSF Aβ1–42 (pg/mL)	561 (256)	604 (283)	594 (277)
Abnormal CSF Aβ1–42 (%)	49	52	51
CSF t-tau (pg/mL)	418 (298)	558 (347)[c]	525 (351)
Abnormal CSF t-tau (%)	51	68	64
CSF p-tau (pg/mL)	64 (33)	80 (47)	77 (45)
Abnormal CSF p-tau (%)	54	71	67
CSF ratio Aβ1–42:t-tau	2.2 (2.2)	1.6 (1.5)	1.8 (1.6)
Outcome at last FU (%)			
No AD	80.8	64.5	72.0
AD	19.2	35.5[b]	28.0
Average FU nondemented subjects (y)	2.35 (0.84)	2.67 (1.31)[c]	2.47 (1.09)
Average time to AD (y)	1.51 (0.71)	2.29 (1.39)[a]	2.07 (1.28)

All volumetric measurements are corrected for intracranial volume. Values are mean (SD).
Key: Aβ1–42, beta amyloid 1–42; AD, Alzheimer's disease; *ApoE*, apolipoprotein E genotype; CSF, cerebrospinal fluid; FU, follow-up; LEAP, learning embeddings for atlas propagation; MMSE, Mini Mental State Examination; MTA-score, qualitative ratings of medial temporal lobe atrophy; p-tau, phosphorylated tau; t-tau, total tau.
[a] $p < 0.01$ for differences between cohorts.
[b] $p < 0.001$ for differences between cohorts.
[c] $p < 0.05$ for differences between cohorts.

3.3. Correlations with AD biomarkers

Manual and LEAP hippocampal volume significantly correlated with CSF t-tau, p-tau, and Aβ/t-tau ratio (all $p < 0.01$; Table 2). The MTA-score correlated with Aβ1–42 ($p < 0.05$) and the ratio Aβ/t-tau ($p < 0.01$). The lateral ventricle volume correlated with Aβ1–42 ($p < 0.01$) and p-tau ($p < 0.05$). Only the LEAP hippocampal volume at baseline was associated with *APOE*-e4 allele status (LEAP volume

Table 2
Correlation with CSF measurements

	Aβ1–42	T-tau	P-tau	Aβ/T-tau
Manual hippocampus	0.09	−0.28[a]	−0.23[a]	0.29[a]
LEAP hippocampus	0.05	−0.32[a]	−0.27[a]	0.33[a]
MTA-score	−0.20[b]	0.11	0.06	−0.24[a]
Lateral ventricle	−0.30[a]	0.12	0.16[b]	−0.10

Shown are Pearson correlation coefficients or Spearman-rank correlation coefficients (correlations with MTA-score).
Key: Aβ, beta amyloid; CSF, cerebrospinal fluid; LEAP, learning embeddings for atlas propagation; MTA-score, qualitative ratings of medial temporal lobe atrophy; P-tau, phosphorylated tau; T-tau, total tau.
[a] $p < 0.01$.
[b] $p < 0.05$.

for *APOE*-e4 carriers 5516 mm^3 vs. 5852 mm^3 for *APOE*-e4 noncarriers; $t = 2.72$; $p < 0.001$).

3.4. Predictors of AD-type dementia in the combined sample

In the total sample, the AUC for prediction of AD-type dementia at the 2-year follow-up was highest for the LEAP hippocampal measurement (0.71) and manual hippocampal measurement (0.71), and it was substantially lower for the MTA-score (0.65) and lateral ventricle (0.60) (Fig. 1). If measurements were dichotomized using the cut point that maximized the Youden index, the overall predictive accuracy for AD after 2 years was best for the manual and LEAP hippocampal volumetric measurements (OR, 6.4–6.5; HR, 4.4–4.5), and lowest for the qualitative rating and lateral ventricle volume (OR, <4; HR, <3) (Table 3). Fig. 2 shows the 2-year dementia risk according to baseline MTL score based on spline analysis. The x-axis shows the degree of atrophy and the y-axis depicts the risk for AD at follow-up associated with the degree of atrophy. Manual and LEAP hippocampal volume had a nonlinear relation with dementia risk indicating that the risk for dementia did not change much with very large or very small volumes. The optimal cut point based on the Youden index for these measures was found halfway through the linear part of the plot.

If measurements were dichotomized using the cut point that provided a sensitivity of 85% similar results were obtained. These analyses also showed that the specificity was higher for manual and LEAP volumetric measurements of hippocampal volume than for MTA-score and lateral ventricle volume.

3.5. Predictors of cognitive decline

Because subjects with MCI might convert to AD-type dementia after the 2-year follow-up, we performed slope analyses with annual cognitive decline up to 5 years as an outcome measure. Subjects with abnormal scores for manual and LEAP hippocampal measurement and MTA-score at baseline declined 2 times as fast on the MMSE ($p < 0.001$) and declined 2.6–4 times as fast on the cognitive composite score ($p < 0.01$) compared with subjects with normal MTL scores (Table 4, Fig. 3). The lateral ventricle volume only predicted decline on the cognitive composite score ($p < 0.01$) (Table 4, Fig. 3).

3.6. Multicenter versus single-center cohort

In the DESCRIPA multicenter cohort, AUC analysis showed best predictive accuracy values for LEAP (0.74) followed by manual hippocampus volume (0.71). In the single-center cohort, manual hippocampus was the best predictor for AD (AUC = 0.69) followed by LEAP (AUC = 0.68) and MTA-score (AUC = 0.65) (Fig. 1). To investigate whether optimal cut points for predicting AD-type dementia differed between the multicenter and single center cohort, we calculated within each cohort the cut point that maximized the Youden index and the cut point that provided a sensitivity of 85% for the prediction of AD-type dementia after 2 year (Supplementary Table 3). These analyses showed that the optimal cut points for the manual and LEAP hippocampal measurement were similar between the multicenter and the single-center cohort with relative differences of less than 2.5% for the cut points based on the Youden index and less than 5.5% for the cut points that provided a sensitivity of 85%.

4. Discussion

The present study showed that volumetric measurements of the MTL are better predictors for AD-type dementia in subjects with MCI than a qualitative rating or the assessment of the lateral ventricle volume.

This is to our knowledge the first study comparing 4 different measures for MTL atrophy in a large sample of MCI subjects from a memory clinic population. Furthermore this is the first study to compare the predictive accuracy of these measurements in single-center and multicenter settings and to investigate the correlations between these measurements and other AD biomarkers.

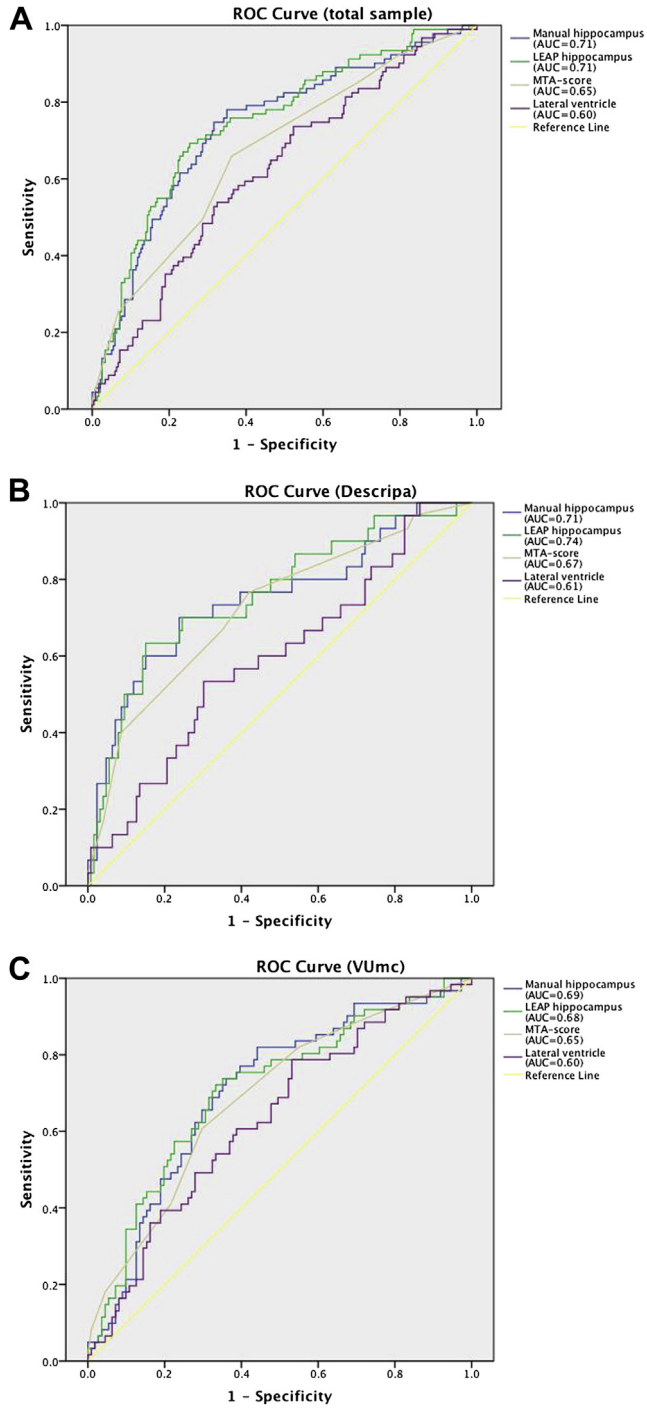

Fig. 1. AUC of a ROC-curve for MRI measurements for (A) total sample, (B) DESCRIPA cohort and (C) VUmc cohort. All volumetric measurements are corrected for intracranial volume. Abbreviations: AUC, area under the curve; LEAP, learning embeddings for atlas propagation; MRI, magnetic resonance imaging; MTA, medial temporal lobe atrophy; ROC, receiver operating characteristics.

Table 3
Predictive accuracy for Alzheimer-type dementia after 2 years

	Cutoff	Sens	Spec	PPV	NPV	OR	HR
Cut point based on Youden index							
Manual hippocampus	7559	0.78	0.65	0.46	0.88	6.47	4.55
LEAP hippocampus	5374	0.66	0.77	0.52	0.85	6.40	4.43
MTA-score	3	0.66	0.64	0.41	0.83	3.40	2.77
Lateral ventricle	58,491	0.53	0.68	0.39	0.79	2.36	2.01
Cut point for a sensitivity of 85%							
Manual hippocampus	8379	0.88	0.37	0.35	0.89	4.22	3.15
LEAP hippocampus	5901	0.87	0.42	0.36	0.89	4.72	3.34
MTA-score	2	0.86	0.30	0.32	0.85	2.57	2.03
Lateral ventricle	34,859	0.84	0.27	0.31	0.81	1.87	2.00

All volumetric measurements are corrected for intracranial volume. Cutoff in mm^3 for volumetric measurements.
Key: HR, hazard ratio; LEAP, learning embeddings for atlas propagation; MTA-score, qualitative ratings of medial temporal lobe atrophy; NPV, negative predictive value; OR, odds ratio; PPV, positive predictive value; Sens, sensitivity; Spec, specificity.

4.1. Comparison of MTL assessment methods

4.1.1. Predictive accuracy

The degree of neurodegeneration in MTL structures is the best MRI marker of imminent conversion to AD, with decreased hippocampal volume being the most robust structural MRI feature (Risacher et al., 2009). In our study, predictive accuracy of both volumetric hippocampal measurements for AD-type dementia was indeed higher than that of the qualitative rating and lateral ventricle measure. This is in line with current evidence stating that manual and automated volumetric methods show similar performance in diagnosing AD (Colliot et al., 2008; Hsu et al., 2002; Lehmann et al., 2010; Sanchez-Benavides et al., 2010a) and that qualitative rating scales or lateral ventricle measurements are less accurate predictors than volumetric methods (Ridha et al., 2007; Urs et al., 2009; Westman et al., 2011).

The slope analyses of cognitive decline over 5 years again yielded similar results with LEAP and manual hippocampal volume at baseline predicting cognitive decline on the MMSE and a cognitive composite score. Slope analysis for the MTA-score predicted cognitive decline equally well as the volumetric measures. This is again in line with previous findings showing that automatically measured volume change in the hippocampus is correlated with decline of performance on the MMSE (Arlt et al., 2012) and that manually measured hippocampal volume reduction is correlated with the severity of impairment on neuropsychological tests (Yavuz et al., 2007). Another study found that performance on the MMSE was directly correlated with hippocampal volume (Laakso et al., 1995).

4.1.2. Relation with other AD biomarkers

Manual and LEAP hippocampal volume significantly correlated with CSF t-tau and p-tau but not with Aβ1–42. The correlation of hippocampal volumetric measurements with CSF tau but not with CSF Aβ is in line with previous studies conducted in subjects with MCI (Apostolova et al., 2010; Carmichael et al., 2012) and subjects with prodromal AD or AD (de Souza et al., 2012). It might be

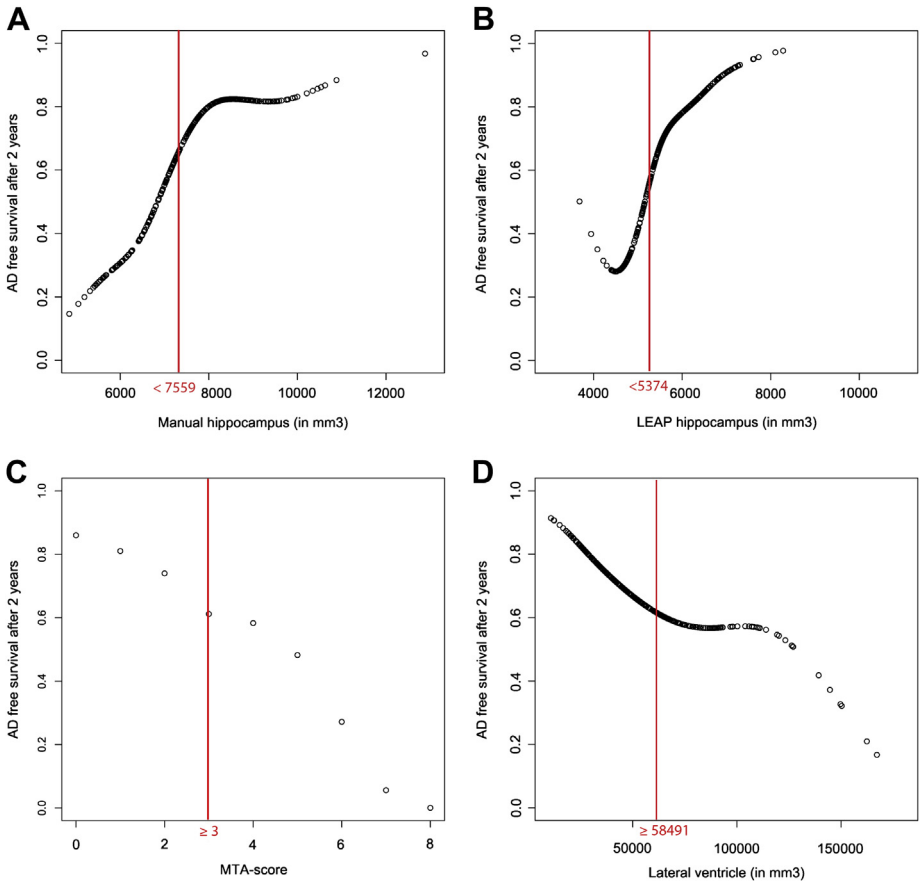

Fig. 2. Spline analyses showing AD-free survival after 2 years as function of MRI measurement in the total cohort for (A) manual hippocampal volume, (B) LEAP hippocampal volume, (C) MTA-score, and (D) lateral ventricle volume. Red line indicates cut point that maximized the Youden index. All volumetric measurements are corrected for intracranial volume. Abbreviations: AD, Alzheimer's disease; LEAP, learning embeddings for atlas propagation; MRI, magnetic resonance imaging; MTA, medial temporal lobe atrophy; MTA-score, qualitative ratings of medial temporal lobe atrophy.

Table 4
Annual cognitive decline over 5 years of follow-up

	MMSE score		Cognitive composite score	
	Baseline score	Slope	Baseline score	Slope
Manual hippocampus				
Normal	27.6 (0.47)[a]	−0.45 (0.09)[a,b]	−0.90 (0.08)	−0.05 (0.02)[c,d]
Abnormal	26.7 (0.47)	−1.04 (0.10)[e]	−1.00 (0.09)	−0.17 (0.03)[e]
LEAP hippocampus				
Normal	27.4 (0.42)[c]	−0.54 (0.09)[a,b]	−0.92 (0.07)	−0.07 (0.02)[b,c]
Abnormal	26.5 (0.46)	−1.13 (0.12)[e]	−1.01 (0.09)	−0.18 (0.04)[e]
MTA-score				
Normal	27.4 (0.43)[f]	−0.49 (0.08)[a,e]	−0.84 (0.07)[c]	−0.05 (0.02)[a,b]
Abnormal	26.9 (0.43)	−1.14 (0.11)[e]	−1.05 (0.08)	−0.20 (0.03)[e]
Lateral ventricle				
Normal	27.4 (0.46)	−0.70 (0.09)[e]	−0.96 (0.07)	−0.01 (0.03)[c]
Abnormal	27.0 (0.48)	−0.94 (0.13)[e]	−1.00 (0.09)	−0.22 (0.05)[c,e]

Data are mean (SD). Slope refers to annual change on the test. A negative slope indicates cognitive decline. Scores dichotomized based on Youden cutoffs.
Key: LEAP, learning embeddings for atlas propagation; MMSE, Mini Mental State Examination; MTA-score, qualitative ratings of medial temporal lobe atrophy.
[a] $p < 0.001$ for baseline score or slope compared with baseline score or slope in the abnormal biomarker group.
[b] $p < 0.01$ for slope different from 0 (this means a statistically significant change over time in test score).
[c] $p < 0.01$ for baseline score or slope compared with baseline score or slope in the abnormal biomarker group.
[d] $p < 0.05$ for slope different from 0 (this means a statistically significant change over time in test score).
[e] $p < 0.001$ for slope different from 0 (this means a statistically significant change over time in test score).
[f] $p < 0.05$ for baseline score or slope compared with baseline score or slope in the abnormal biomarker group.

explained by the observation that antemortem hippocampal volume significantly correlated with the density of neurofibrillary tangles at autopsy (Csernansky et al., 2004; Jack et al., 2002) but not with Aβ plaque load (Csernansky et al., 2004). The qualitative MTA-score and lateral ventricle volume correlated with CSF Aβ1–42. This correlation might indicate that these MTL measures in part reflect generalized brain atrophy because previous studies showed that lower Aβ1–42 levels but not t-tau levels were associated with total brain atrophy and ventricular volume (Wahlund and Blennow, 2003).

4.1.3. Overlap between measurements

Although the LEAP and manual hippocampal measurement correlated highly ($r = 0.71$) and scores showed 80% overlap, the ICC of 0.38 indicated a moderate agreement. This might be because the LEAP volume was consistently lower, because the ICC is sensitive for absolute sizes. A previous study that compared LEAP with manual measurements found a much higher ICC of 0.89 (Wolz et al., 2010a). Although in this study volume obtained using LEAP was systematically lower than volumes obtained using the manual volumetric measurement, this difference was smaller than in our study. Differences in absolute volumes might be because of the use of different borders for hippocampal delineation. Major parts of the hippocampus are included in both volumetric methods with the only difference being that the LEAP method misses some of the hippocampal head, some of the alveus, and some of the fimbria, and ends a few slices earlier than manual hippocampal outlining. Despite the differences in raw volume, both methods have a comparable diagnostic accuracy indicating that these parts are less important for AD pathology. It is also likely that the cutoff points for each hippocampal method likely reflected a similar degree of abnormality. Namely, if we defined the cutoff that provided a sensitivity of 0.85, the specificity of each method was equal. When the cutoff was used that optimized the Youden index, small differences in sensitivity and specificity were found but the combination of sensitivity and specificity expressed as odds ratio was the same. Differences in absolute volumes between different measurement protocols might be reduced in the future as efforts are made toward the harmonization of an MRI segmentation protocol for hippocampal delineation (Boccardi et al., 2011; Frisoni and Jack, 2011). Because manual volumetry is used as the standard against which automated segmentation algorithms are assessed, future synchronization and comparison of both techniques will be facilitated. Correlations between manual and LEAP volumetric measurements and MTA-score were low with a moderate overlap between both methods. The same pattern was found for correlations and overlap between both volumetric hippocampal measurements and expansion of the lateral ventricle.

Our findings indicate, in line with previous studies, that volumetric and qualitative measures of MTL atrophy measure different aspects of AD pathology (Visser et al., 2002b; Wahlund et al., 1999). Manual and LEAP hippocampal volume correlated with each other, both showed high predictive accuracy values, and both correlated with t-tau and p-tau but not with Aβ1–42. In contrast, visually rated MTA-score and expansion of the lateral ventricle correlated highly with each other but not with the volumetric measures. They showed lower predictive accuracy for AD conversion than the volumetric measures and correlated highly with Aβ in CSF rather than CSF tau. As already discussed in this report, it is possible that both the MTA-score and lateral ventricle volume reflect a widening of ventricular spaces (i.e., the temporal horn), which might be indicative of generalized atrophy rather than atrophy of the hippocampus alone.

4.2. Comparison of cohorts

None of the MTL measures showed major differences in predictive accuracy between the single and multicenter cohort. When optimal cut points for each measure were calculated for each cohort separately these cut points only slightly differed between the cohorts for the volumetric hippocampal measurements, which shows that volumetric measurements display stable cut points across different cohorts. Some differences were noted between the cohorts on cut points for the MTA-score and lateral ventricle, suggesting that MTA-score and lateral ventricle are more sensitive for cohort differences.

4.3. Technical considerations

In general automated measurements are more susceptible to scanner and scan protocol variability. In our study, no significant differences between both methods were found with LEAP showing similar or even slightly better performance than manual volumetry. This suggests that automated measurements can be performed in multicenter studies without strict standardization of scan protocols. The percentage of technical failures was 0.07% manual hippocampal measurement and 0.03% for LEAP hippocampal measurement. All failures were observed in 2 DESCRIPA sites that apparently used scan frequencies which occasionally led to technical errors during the preprocessing phase (e.g., intensity problems), making these particular scans unsuitable for applying subsequent volumetric measurements. SIENAX-related measurements (ICV and lateral ventricular volume) yielded a higher technical failure rate (1% for lateral ventricle) and were observed across all sites. These were because of calculation errors or a combination of software and image quality problems. These findings indicate that the lateral ventricle measurement is more sensitive for specific differences in image quality than a manual or automated hippocampal measurement. It should be noted than scan sequences were designed for routine clinical practice and not for automated

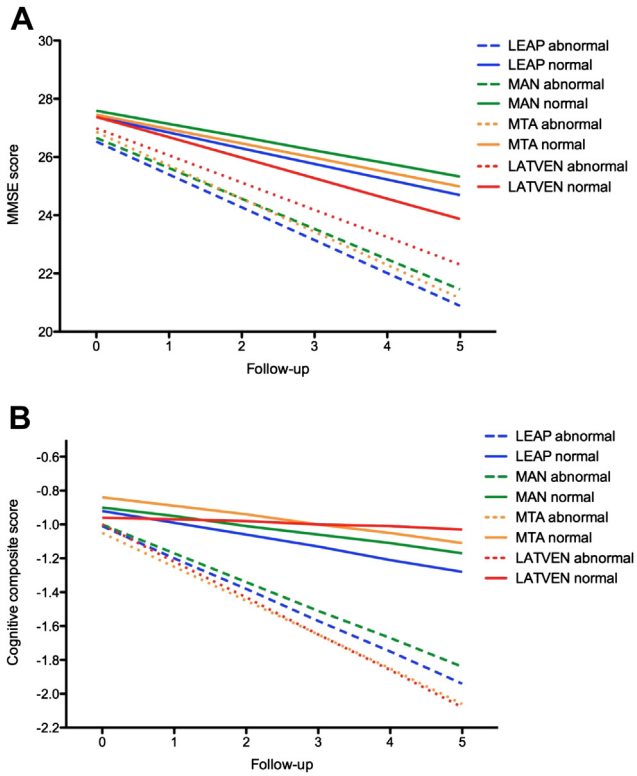

Fig. 3. Slope analyses showing annual cognitive decline on (A) the MMSE and (B) a cognitive composite score as predicted by baseline medial temporal lobe measurements. Abnormal MRI values at baseline were defined as manual hippocampal volume <7559 mm^3, LEAP hippocampal volume <5374 mm^3, MTA-score ≥3, and lateral ventricle volume ≥58491 mm^3. The solid lines indicate the subjects with normal values. The dotted lines indicates subjects with abnormal values. All volumetric measurements are corrected for intracranial volume. The cognitive composite score = average z-score of learning or delayed recall of a word list learning test or equivalent memory test, the Trail Making Test part A, Trail Making Test part B, verbal fluency, Rey figure copy test, or an equivalent test. Abbreviations: LATVEN, lateral ventricle volume; LEAP, learning embeddings for atlas propagation; MAN, manual hippocampal volume; MMSE, Mini Mental State Examination; MRI, magnetic resonance imaging; MTA, medial temporal lobe atrophy score; MTA-score, qualitative ratings of medial temporal lobe atrophy.

measurements and it is likely that failure rate could be reduced using protocols optimized for automated measurements.

4.4. Limitations

The present study had several limitations. There were differences in subject characteristics between the single-center and multicenter cohort. Subjects from the single-center VUmc cohort were more severely impaired and had lower baseline MMSE scores, more severe memory impairment, a lower hippocampal volume, higher CSF tau, and a higher conversion rate to dementia. Despite these differences, however, the predictive accuracy was similar in each cohort. Another limitation is that the cutoffs of all MRI measurements were determined within a study population that also included the subjects from the present analyses. This could have led to an overestimation of the predictive accuracy. However, it is unlikely that it influenced our findings with respect to the differences in predictive accuracy between MRI measurements, because we used the same method to define the cut point for each measurement. Follow-up data with AD diagnosis for all subjects was only available for a relatively short follow-up interval (2 years). For clinical trials, however, short-term prognosis might be important. For subjects in whom long-term follow-up data were available, predictive accuracy for annual cognitive decline up to 5 years was additionally investigated. The diagnosis of AD at follow-up was not validated neuropathologically which might have possibly led to the misclassification of some cases. We used scanners with different field strengths (1T and 1.5T) which reflected real-life situations in which scanners and magnetic field strengths vary. However, these differences might have introduced bias. We therefore compared the hippocampal volume between subjects scanned on a 1.0T scanner ($n = 127$) and a 1.5T scanner ($n = 201$) after correction for age, sex, educational level, baseline MMSE score, and follow-up diagnosis. The difference in volume between 1.0T and 1.5T scanners was 0.2% for the LEAP method ($p = 0.9$; van Rossum et al., 2012), 0.9% for manual hippocampal volume ($p = 0.6$), and 1.5% for the lateral ventricle volume ($p = 0.8$). We also tested whether field strength modulated the effect of the volumetric measures on conversion to AD-type dementia. For none of the measures the interaction between field strength and volumetric measure was statistically significantly associated with conversion ($p > 0.15$). A recent study also found a limited effect of field strength on hippocampal volume. This study compared the hippocampal volume measured by the LEAP method between subjects scanned both on a 1.5T and 3T scanner and found a very high correlation between the measurements on each scanner ($r = 0.98$) (Wolz 2013, unpublished data). The volumes measured on 3T were on average 24.4 mm^3 or 1.17% larger than on 1.5T (Wolz 2013, unpublished). This variability was similar to that for volumes rescanned on scanners with the same field strength (1.5%) (Wolz 2013, unpublished). Taken together, it is unlikely that the small difference in field strength in our study had a major effect on the volumetric measurements in our study. Different methods were used for ICV correction of both hippocampal measurements (see section 2.3.2.). Both correction methods were thoroughly compared and showed high correlations (ICC = 0.93 for LEAP with MNI vs. FSL scaling and 0.95 for manual volumetric measurement with MNI vs. FSL scaling) and similar results regarding predictive accuracy. As a result of this analysis, ICV correction of each individual method was applied. It can be considered a strength that for this diagnostic study a population from a memory clinic setting was used. A consequence however is that these findings might not be applicable to other settings, including the general population.

4.5. Clinical implications

Volumetric measurements of the MTL are the best predictors for AD-type dementia in subjects with MCI. Both volumetric measurements strongly correlate with CSF markers of neuronal injury (CSF t-tau and p-tau), are able to predict cognitive decline, and show consistent cutoff values between different cohorts. LEAP hippocampal volume has the advantage over manual volumetry in that it needs much less rater time and shows no interrater variability effects. In addition, LEAP has a low technical failure rate. Visual rating scales are also quick and easy to perform but show lower predictive accuracy rates and higher inter- and intra-individual variability effects (DeCarli et al., 1990) compared with LEAP volumetric measurement. Another disadvantage of visual rating scales, although outside the scope of this study, is that they cannot detect subtle atrophy progression and are thus insensitive to change over time (Ridha et al., 2007; Urs et al., 2009). Because the hippocampus is among the first areas affected by the disease (Chupin et al., 2009; Lötjönen et al., 2011), repeated measurement of its volume is clinically important. LEAP hippocampal measurement is suitable for implementation in clinical practice with on average 4-minute control time on a standard computer. Cut points that maximized the balance between sensitivity and specificity as expressed by the Youden index or that provided a sensitivity of 85%

were defined. The cut point based on the Youden index might be preferred because it has shown to be more consistent between the multi- and single-center cohort (Supplementary Table 3). Depending on the clinical needs, other cut points can be chosen, that for example maximize the positive or negative predictive value (Bartlett et al., 2012).

Any choice for a specific cut point has a trade-off between the chance of missing the disease (false negative rate, 1-sensitivity) or incorrectly diagnosing someone as having the disease (false positive rate, 1-specificity). The choice will therefore depend on its clinical use. For example, if a treatment for MCI caused by AD would be available the choice might depend on the safety profile of the treatment. If a treatment has many or severe side effects, high specificity is more important than a high sensitivity and the opposite is true for treatments with few side effects. If treatment is not available and biomarkers are used for diagnosis, one might prefer to use a cut point with a low false positive rate because an incorrect diagnosis of AD might have a major negative effect on the patient.

4.6. Future directions

Future MRI studies need to investigate abnormalities in AD signature regions in and outside the MTL. A recent study found that subjects with future cognitive impairment (preclinical AD and MCI) also showed reduced brain volume in posterior cingulate and/or precuneus and orbitofrontal cortex, at least 4 years before any cognitive symptoms (Tondelli et al., 2012). Other structural MRI studies also found abnormalities in AD or MCI outside the MTL region such as the corpus callosum (Chen et al., 2009; Di Paola et al., 2010; Serra et al., 2010; Wang and Su, 2006), cingulum (Callen et al., 2001; Choo et al., 2010; Jones et al., 2006), parietal (Jacobs et al., 2012), temporal lobe other than MTL (Chincarini et al., 2011), and frontal lobe (Burgmans et al., 2009). Future MRI studies need to investigate abnormalities in AD signature regions in and outside the MTL.

4.7. Conclusion

Volumetric hippocampal measurements are the best predictors of conversion to AD-type dementia in subjects with MCI after 2-year follow-up and are able to predict annual cognitive decline. Because of the limited rater time, LEAP automated hippocampal measurement might be preferred.

Disclosure statement

Leah Burns is an employee of Bristol-Myers Squibb. Philip Scheltens serves or has served on the advisory boards of: Genentech, Novartis, Roche, Danone, Nutricia, Baxter, and Lundbeck. He has been a speaker at symposia organized by Lundbeck, Merz, Danone, Novartis, Roche, and Genentech. For all his activities he receives no personal compensation. He serves on the editorial board of *Alzheimer's Research & Therapy* and *Alzheimer Disease & Associated Disorders*, is a member of the scientific advisory board of the EU Joint Programming Initiative and the French National Plan Alzheimer. The Alzheimer Center receives unrestricted funding from various sources through the VUmc Fonds. Pablo Lapuerta is a former employee of Bristol-Myers Squibb and current CMO of Lexicon Pharmaceuticals. Frederik Barkhof has received the following funding: consulting fees or honoraria from UCB, Bayer Schering, Sanofi-Aventis, Novartis, Roche, Merck-Serono, Synthon BV, Janssen Research, Lundbeck, and Biogen-Idec, and is a member of UCB, Bayer Schering, Sanofi-Aventis, Novartis, and Roche advisory boards. Robin Wolz is part-time consultant at IXICO. Daniel Rueckert is cofounder and consultant at IXICO. Magdalini Tsolaki serves on scientific advisory boards for Novartis and Pfizer. Flavio Nobili served on the European Bayer Florbetaben board in October 2010 and in the Italian Bayer Florbetaben board in 2011. Lennart Minthon serves on advisory boards for Pfizer, Sweden. Lutz Frölich has received funding from pharmaceutical companies involved in the manufacture and marketing of drugs or medicinal products for Alzheimer's disease including: Avid-Eli Lilly, Astra-Zeneca, Baxter, Bayer, Eisai, GE Health Care, Janssen-Cilag, Lundbeck, MerckSharpe & Dohme, Merz Pharma, Novartis, Pfizer, and Schering-Plough. Harald Hampel has received lecture honoraria and/or research grants and/or travel funding and/or participated in scientific advisory boards of pharmaceutical companies involved in the manufacture and marketing of diagnostics and/or drugs or medicinal products for Alzheimer's disease including Boehringer-Ingelheim, BRAHMS, BMS, ELAN, Wyeth, Novartis, Eisai, Pfizer, Schwabe, Sanofi-Aventis, Roche, GE, Astra-Zeneca, Avid-Eli Lilly and Company, Janssen-Cilag, and Merz Pharmaceuticals. Pieter Jelle Visser has served as an advisory board member of Myriad, Guidage study Ipsen, and Bristol-Myers Squibb. He receives or has received research grants from Bristol-Myers Squibb, European Commission 6th and 7th Framework programme, Life Sciences, Genomics and Biotechnology for Health, Diagenic, Norway, and Innogenetics, Belgium. The remaining authors report no disclosures. The medical ethics committee at each center approved the study. All patients provided informed consent.

Acknowledgements

Lies Clerx receives research support from the Center for Translational Molecular Medicine, project LeARN (grant 02N-01). Ineke A. van Rossum received research support via Lipididiet from European Community's Seventh Framework Programme (FP7/2007-2013), grant agreement no 211696. Bristol-Myers Squibb (BMS) financially contributed to the MRI analyses. Statistical analysis and writing of the manuscript was performed within the framework of CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project LeARN (grant 02N-01).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.neurobiolaging.2013.02.002.

References

Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D., 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. Neuroimage 46, 726–738.

American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders, fourth ed. American Psychiatric Association, Washington, D.C.

Andreasen, N., Hesse, C., Davidsson, P., Minthon, L., Wallin, A., Winblad, B., Vanderstichele, H., Vanmechelen, E., Blennow, K., 1999. Cerebrospinal fluid beta-amyloid(1-42) in Alzheimer disease: differences between early- and late-onset Alzheimer disease and stability during the course of disease. Arch. Neurol. 56, 673–680.

Apostolova, L.G., Dinov, I.D., Dutton, R.A., Hayashi, K.M., Toga, A.W., Cummings, J.L., Thompson, P.M., 2006. 3D comparison of hippocampal atrophy in amnestic mild cognitive impairment and Alzheimer's disease [erratum in 2007;130:2474]. Brain 129, 2867–2873.

Apostolova, L.G., Hwang, K.S., Andrawis, J.P., Green, A.E., Babakchanian, S., Morra, J.H., Cummings, J.L., Toga, A.W., Trojanowski, J.Q., Shaw, L.M., Jack Jr., C.R., Petersen, R.C., Aisen, P.S., Jagust, W.J., Koeppe, R.A., Mathis, C.A., Weiner, M.W., Thompson, P.M., Alzheimer's Disease Neuroimaging Initiative, 2010. 3D PIB and CSF biomarker associations with hippocampal atrophy in ADNI subjects. Neurobiol. Aging 31, 1284–1303.

Arlt, S., Buchert, R., Spies, L., Eichenlaub, M., Lehmbeck, J.T., Jahn, H., 2012. Association between fully automated MRI-based volumetry of different brain regions and neuropsychological test performance in patients with amnestic mild

cognitive impairment and Alzheimer's disease. Eur. Arch. Psychiatry Clin. Neurosci.
Barnes, J., Ourselin, S., Fox, N.C., 2009. Clinical application of measurement of hippocampal atrophy in degenerative dementias. Hippocampus 19, 510–516.
Bartlett, J.W., Frost, C., Mattsson, N., Skillbäck, T., Blennow, K., Zetterberg, H., Schott, J.M., 2012. Determining cut-points for Alzheimer's disease biomarkers: statistical issues, methods and challenges. Biomark. Med. 6, 391–400.
Blennow, K., Wallin, A., Ågren, H., Spenger, C., Siegfried, J., Vanmechelen, E., 1995. Tau protein in cerebrospinal fluid: a biochemical diagnostic marker for axonal degeneration in Alzheimer's disease? Mol. Chem. Neuropathol. 26, 231–245.
Boccardi, M., Ganzola, R., Bocchetta, M., Pievani, M., Redolfi, A., Bartzokis, G., Camicioli, R., Csernansky, J.G., de Leon, M.J., deToledo-Morrell, L., Killiany, R.J., Lehéricy, S., Pantel, J., Pruessner, J.C., Soininen, H., Watson, C., Duchesne, S., Jack Jr., C.R., Frisoni, G.B., 2011. Survey of protocols for the manual segmentation of the hippocampus: preparatory steps towards a joint EADC-ADNI harmonized protocol. J. Alzheimers Dis. 26 (suppl. 3), 61–75.
Burgmans, S., van Boxtel, M.P., Smeets, F., Vuurman, E.F., Gronenschild, E.H., Verhey, F.R., Uylings, H.B., Jolles, J., 2009. Prefrontal cortex atrophy predicts dementia over a six-year period. Neurobiol. Aging 30, 1413–1419.
Callen, D.J., Black, S.E., Gao, F., Caldwell, C.B., Szalai, J.P., 2001. Beyond the hippocampus: MRI volumetry confirms widespread limbic atrophy in AD. Neurology 57, 1669–1674.
Carmichael, O., Xie, J., Fletcher, E., Singh, B., DeCarli, C., 2012. Localized hippocampus measures are associated with Alzheimer pathology and cognition independent of total hippocampal volume. Neurobiol. Aging 33, 1124.e1131–1124.e1141.
Carmichael, O.T., Aizenstein, H.A., Davis, S.W., Becker, J.T., Thompson, P.M., Meltzer, C.C., Liu, Y., 2005. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. Neuroimage 27, 979–990.
Chen, T.F., Lin, C.C., Chen, Y.F., Liu, H.M., Hua, M.S., Huang, Y.C., Chiu, M.J., 2009. Diffusion tensor changes in patients with amnesic mild cognitive impairment and various dementias. Psychiatry Res. 173, 15–21.
Chincarini, A., Bosco, P., Calvini, P., Gemme, G., Esposito, M., Olivieri, C., Rei, L., Squarcia, S., Rodriguez, G., Bellotti, R., Cerello, P., De Mitri, I., Retico, A., Nobili, F., Alzheimer's Disease Neuroimaging Initiative, 2011. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. Neuroimage 58, 469–480.
Choo, I.L., Lee, D.Y., Oh, J.S., Lee, J.S., Lee, D.S., Song, I.C., Youn, J.C., Kim, S.G., Kim, K.W., Jhoo, J.H., Woo, J.I., 2010. Posterior cingulate cortex atrophy and regional cingulum disruption in mild cognitive impairment and Alzheimer's disease. Neurobiol. Aging 31, 772–779.
Chou, Y.Y., Lepore, N., Saharan, P., Madsen, S.K., Hua, X., Jack, C.R., Shaw, L.M., Trojanowski, J.Q., Weiner, M.W., Toga, A.W., Thompson, P.M., Alzheimer's Disease Neuroimaging Initiative, 2010. Ventricular maps in 804 ADNI subjects: correlations with CSF biomarkers and clinical decline. Neurobiol. Aging 31, 1386–1400.
Chupin, M., Gérardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehéricy, S., Benali, H., Garnero, L., Colliot, O., Alzheimer's Disease Neuroimaging Initiative, 2009. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. Hippocampus 19, 579–587.
Colliot, O., Chételat, G., Chupin, M., Desgranges, B., Magnin, B., Benali, H., Dubois, B., Garnero, L., Eustache, F., Lehéricy, S., 2008. Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. Radiology 248, 194–201.
Csernansky, J.G., Hamstra, J., Wang, L., McKeel, D., Price, J.L., Gado, M., Morris, J.C., 2004. Correlations between antemortem hippocampal volume and postmortem neuropathology in AD subjects. Alzheimer Dis. Assoc. Disord. 18, 190–195.
DeCarli, C., Frisoni, G.B., Clark, C.M., Harvey, D., Grundman, M., Petersen, R.C., Thal, L.J., Jin, S., Jack Jr., C.R., Scheltens, P., Alzheimer's Disease Cooperative Study Group, 2007. Qualitative estimates of medial temporal atrophy as a predictor of progression from mild cognitive impairment to dementia. Arch. Neurol. 64, 108–115.
DeCarli, C.M., Kaye, J.A., Horwitz, B.P., Rapoport, S.I., 1990. Critical analysis of the use of computer-assisted transverse axial tomography to study human brain in aging and dementia of the Alzheimer type. Neurology 40, 872–883.
de Souza, L.C., Chupin, M., Lamari, F., Jardel, C., Leclercq, D., Colliot, O., Lehéricy, S., Dubois, B., Sarazin, M., 2012. CSF tau markers are correlated with hippocampal volume in Alzheimer's disease. Neurobiol. Aging 33, 1253–1257.
Di Paola, M., Di Iulio, F., Cherubini, A., Blundo, C., Casini, A.R., Sancesario, G., Passafiume, D., Caltagirone, C., Spalletta, G., 2010. When, where, and how the corpus callosum changes in MCI and AD: a multimodal MRI study. Neurology 74, 1136–1142.
Echávarri, C., Aalten, P., Uylings, H.B., Jacobs, H., Visser, P., Gronenschild, E., Verhey, F.R., Burgmans, S., 2011. Atrophy in the parahippocampal gyrus as an early biomarker of Alzheimer's disease. Brain Struct. Funct. 215, 265–271.
Frisoni, G.B., Jack, C.R., 2011. Harmonization of magnetic resonance-based manual hippocampal segmentation: a mandatory step for wide clinical use. Alzheimers Dement. 7, 171–174.
Galton, C.J., Gomez-Anson, B., Antoun, N., Scheltens, P., Patterson, K., Graves, M., Sahakian, B.J., Hodges, J.R., 2001. Temporal lobe rating scale: application to Alzheimer's disease and frontotemporal dementia. J. Neurol. Neurosurg. Psychiatry 70, 165–173.
Hanley, J.A., McNeil, B.J., 1983. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 148, 839–843.
Heagerty, P.J., Lumley, T., Pepe, M.S., 2000. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 56, 337–344.
Henneman, W.J., Sluimer, J.D., Cordonnier, C., Baak, M.M., Scheltens, P., Barkhof, F., van der Flier, W.M., 2009. MRI biomarkers of vascular damage and atrophy predicting mortality in a memory clinic population. Stroke 40, 492–498.
Hsu, Y.Y., Schuff, N., Du, A.T., Mark, K., Zhu, X., Hardin, D., Weiner, M.W., 2002. Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia. J. Magn. Reson. Imaging 16, 305–310.
Jack, C.R., 1994. MRI-based hippocampal volume measurements in epilepsy. Epilepsia 35, S21–S29.
Jack Jr., C.R., Dickson, D.W., Parisi, J.E., Xu, Y.C., Cha, R.H., O'Brien, P.C., Edland, S.D., Smith, G.E., Boeve, B.F., Tangalos, E.G., Kokmen, E., Petersen, R.C., 2002. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. Neurology 58, 750–757.
Jack Jr., C.R., Shiung, M.M., Gunter, J.L., O'Brien, P.C., Weigand, S.D., Knopman, D.S., Boeve, B.F., Ivnik, R.J., Smith, G.E., Cha, R.H., Tangalos, E.G., Petersen, R.C., 2004. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. Neurology 62, 591–600.
Jacobs, H.I., Van Boxtel, M.P., Jolles, J., Verhey, F.R., Uylings, H.B., 2012. Parietal cortex matters in Alzheimer's disease: an overview of structural, functional and metabolic findings. Neurosci. Biobehav. Rev. 36, 297–309.
Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841.
Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156.
Jones, B.F., Barnes, J., Uylings, H.B., Fox, N.C., Frost, C., Witter, M.P., Scheltens, P., 2006. Differential regional atrophy of the cingulate gyrus in Alzheimer disease: a volumetric MRI study. Cereb. Cortex 16, 1701–1708.
Kolen, M.J., Brennan, R.L., 1995. Test Equating: Methods and Practices. Springer, New York.
Laakso, M.P., Soininen, H., Partanen, K., Helkala, E., Hartikainen, P., Vaino, P., Hallikainen, M., Hänninen, T., Riekkinen Sr., P.J., 1995. Volumes of hippocampus, amygdala and frontal lobes in the MRI-based diagnosis of early Alzheimer's disease: correlation with memory functions. J. Neural Transm. Park. Dis. Dement. Sect. 9, 73–86.
Lehmann, M., Douiri, A., Kim, L.G., Modat, M., Chan, D., Ourselin, S., Barnes, J., Fox, N.C., 2010. Atrophy patterns in Alzheimer's disease and semantic dementia: a comparison of FreeSurfer and manual volumetric measurements. Neuroimage 49, 2264–2274.
Levy-Cooperman, N., Ramirez, J., Lobaugh, N.J., Black, S.E., 2008. Misclassified tissue volumes in Alzheimer disease patients with white matter hyperintensities. Stroke 39, 1134–1141.
Lötjönen, J., Wolz, R., Koikkalainen, J., Julkunen, V., Thurfjell, L., Lundqvist, R., Waldemar, G., Soininen, H., Rueckert, D., Alzheimer's Disease Neuroimaging Initiative, 2011. Fast and robust extraction of hippocampus from MR images for diagnostics of Alzheimer's disease. Neuroimage 56, 185–196.
Masters, C.L., Beyreuther, K., 2006. Alzheimer's centennial legacy: prospects for rational therapeutic intervention targeting the Abeta amyloid pathway. Brain 129, 2823–2839.
McHugh, T.L., Saykin, A.J., Wishart, H.A., Flashman, L.A., Cleavinger, H.B., Rabin, L.A., Mamourian, A.C., Shen, L., 2007. Hippocampal volume and shape analysis in an older adult population. Clin. Neuropsychol. 21, 130–145.
McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M., 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 34, 939–944.
Morris, J.C., 1993. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 43, 2412–2414.
Mulder, C., Verwey, N.A., van der Flier, W.M., Bouwman, F.H., Kok, A., van Elk, E.J., Scheltens, P., Blankenstein, M.A., 2010. Amyloid-β(1-42), total tau, and phosphorylated tau as cerebrospinal fluid biomarkers for the diagnosis of Alzheimer disease. Clin. Chem. 56, 248–253.
Nestor, S.M., Rupsingh, R., Borrie, M., Smith, M., Accomazzi, V., Wells, J.L., Fogarty, J., Bartha, R., Alzheimer's Disease Neuroimaging Initiative, 2008. Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. Brain 131, 2443–2454.
Petersen, R.C., 2004. Mild cognitive impairment as a diagnostic entity. J. Int. Med. 256, 183–194.
Petersen, R.C., Smith, G.E., Waring, S.C., Ivnik, R.J., Tangalos, E.G., Kokmen, E., 1999. Mild cognitive impairment: clinical characterization and outcome. Arch. Neurol. 56, 303–308.
Ridha, B.H., Barnes, J., van de Pol, L.A., Schott, J.M., Boyes, R.G., Siddique, M.M., Rossor, M.N., Scheltens, P., Fox, N.C., 2007. Application of automated medial temporal lobe atrophy scale to Alzheimer disease. Arch. Neurol. 64, 849–854.
Risacher, S.L., Saykin, A.J., West, J.D., Shen, L., Firpi, H.A., McDonald, B.C., Alzheimers Disease Neuroimaging Initiative (ADNI), 2009. Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. Curr. Alzheimer Res. 6, 347–361.
Rusinek, H.P., Endo, Y.M., De Santi, S.P., Frid, D.B., Tsui, W.H., Segal, S.M., Convit, A., de Leon, M.J., 2004. Atrophy rate in medial temporal lobe during progression of Alzheimer disease. Neurology 63, 2354–2359.
Sanchez-Benavides, G., Gomez-Anson, B., Molinuevo, J.L., Blesa, R., Monte, G.C., Buschke, H., Peña-Casanova, J., 2010a. Medial temporal lobe correlates of

memory screening measures in normal aging, MCI, and AD. J. Geriatr. Psychiatry Neurol. 23, 100–108.
Sanchez-Benavides, G., Gomez-Ansun, B., Sainz, A., Vives, Y., Delfino, M., Peoa-Casanova, J., 2010b. Manual validation of FreeSurfer's automated hippocampal segmentation in normal aging, mild cognitive impairment, and Alzheimer disease subjects. Psychiatry Res. 181, 219–225.
Scheltens, P., Leys, D., Barkhof, F., Huglo, D., Weinstein, H.C., Vermersch, P., Kuiper, M., Steinling, M., Wolters, E.C., Valk, J., 1992. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. J. Neurol. Neurosurg. Psychiatry 55, 967–972.
Schoonenboom, N.S., van der Flier, W.M., Blankenstein, M.A., Bouwman, F.H., Van Kamp, G.J., Barkhof, F., Scheltens, P., 2008. CSF and MRI markers independently contribute to the diagnosis of Alzheimer's disease. Neurobiol. Aging 29, 669–675.
Serra, L., Cercignani, M., Lenzi, D., Perri, R., Fadda, L., Caltagirone, C., Macaluso, E., Bozzali, M., 2010. Grey and white matter changes at different stages of Alzheimer's disease. J. Alzheimers Dis. 19, 147–159.
Shen, L., Saykin, A., Kim, S., Firpi, H., West, J., Risacher, S., McDonald, B.C., McHugh, T.L., Wishart, H.A., Flashman, L.A., 2010. Comparison of manual and automated determination of hippocampal volumes in MCI and early AD. Brain Imaging Behav. 4, 86–95.
Sluimer, J.D., van der Flier, W.M., Karas, G.B., Fox, N.C., Scheltens, P., Barkhof, F., Vrenken, H., 2008. Whole-brain atrophy rate and cognitive decline: longitudinal MR study of memory clinic patients. Radiology 248, 590–598.
Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155.
Smith, S.M., De Stefano, N., Jenkinson, M., Matthews, P.M., 2001. Normalized accurate measurement of longitudinal brain change. J. Comp. Assist. Tomogr. 25, 466–475.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23, S208–S219.
Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, A., De Stefano, N., 2002. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage 17, 479–489.
Tondelli, M., Wilcock, G.K., Nichelli, P., De Jager, C.A., Jenkinson, M., Zamboni, G., 2012. Structural MRI changes detectable up to ten years before clinical Alzheimer's disease. Neurobiol. Aging 33, 825.e825–825.e836.
Urs, R., Potter, E., Barker, W., Appel, J., Loewenstein, D.A., Zhao, W., Duara, R., 2009. Visual rating system for assessing magnetic resonance images: a tool in the diagnosis of mild cognitive impairment and Alzheimer disease. J. Comput. Assist. Tomogr. 33, 73–78.
van de Pol, L.A., Barnes, J., Scahill, R.I., Frost, C., Lewis, E.B., Boyes, R.G., van Schijndel, R.A., Scheltens, P., Fox, N.C., Barkhof, F., 2007a. Improved reliability of hippocampal atrophy rate measurement in mild cognitive impairment using fluid registration. Neuroimage 34, 1036–1041.
van de Pol, L.A., van der Flier, W.M., Korf, E.S., Fox, N.C., Barkhof, F., Scheltens, P., 2007b. Baseline predictors of rates of hippocampal atrophy in mild cognitive impairment. Neurology 69, 1491–1497.
van de Pol, L.A., Verhey, F., Frisoni, G.B., Tsolaki, M., Papapostolou, P., Nobili, F., Wahlund, L.O., Minthon, L., Frölich, L., Hampel, H., Soininen, H., Knol, D.L., Barkhof, F., Scheltens, P., Visser, P.J., 2009. White matter hyperintensities and medial temporal lobe atrophy in clinical subtypes of mild cognitive impairment: the DESCRIPA study. J. Neurol. Neurosurg. Psychiatry 80, 1069–1074.
van Rossum, I.A., Vos, S.J., Burns, L., Knol, D.L., Scheltens, P., Soininen, H., Wahlund, L.O., Hampel, H., Tsolaki, M., Minthon, L., L'italien, G., van der Flier, W.M., Teunissen, C.E., Blennow, K., Barkhof, F., Rueckert, D., Wolz, R., Verhey, F., Visser, P.J., 2012. Injury markers predict time to dementia in subjects with MCI and amyloid pathology. Neurology 79, 1801–1816.
Varela-Nallar, L., Aranguiz, F.C., Abbott, A.C., Slater, P.G., Inestrosa, N.C., 2010. Adult hippocampal neurogenesis in aging and Alzheimer's disease. Birth Defects Res. C Embryo Today 90, 284–296.
Vellas, B., Andrieu, S., Sampaio, C., Wilcock, G., 2007. Disease-modifying trials in Alzheimer's disease: a European task force consensus. Lancet Neurol. 6, 56–62.
Visser, P.J., Scheltens, P., Verhey, F.R., Schmand, B., Launer, L., Jolles, J., Jonker, C., 1999. Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment. J. Neurol. 246, 477–485.
Visser, P.J., Verhey, F., Knol, D.L., Scheltens, P., Wahlund, L.O., Freund-Levi, Y., Tsolaki, M., Minthon, L., Wallin, A.K., Hampel, H., Bürger, K., Pirttila, T., Soininen, H., Rikkert, M.O., Verbeek, M.M., Spiru, L., Blennow, K., 2009. Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: a prospective cohort study. Lancet Neurol. 8, 619–627.
Visser, P.J., Verhey, F.R., Boada, M., Bullock, R., De Deyn, P.P., Frisoni, G.B., Frolich, L., Hampel, H., Jolles, J., Jones, R., Minthon, L., Nobili, F., Olde Rikkert, M., Ousset, P.J., Rigaud, A.S., Scheltens, P., Soininen, H., Spiru, L., Touchon, J., Tsolaki, M., Vellas, B., Wahlund, L.O., Wilcock, G., Winblad, B., 2008. Development of screening guidelines and clinical criteria for predementia Alzheimer's disease. Neuroepidemiology 30, 254–265.
Visser, P.J., Verhey, F.R., Hofman, P.A., Jolles, J., 2002a. Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. J. Neurol. Neurosurg. Psychiatry 72, 491–497.
Visser, P.J., Verhey, F.R., Hofman, P.A., Scheltens, P., Jolles, J., 2002b. Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. J. Neurol. Neurosurg. Psychiatry 72, 491–497.
Vos, S., van Rossum, I., Burns, L., Knol, D., Scheltens, P., Soininen, H., Wahlund, L.O., Hampel, H., Tsolaki, M., Minthon, L., Handels, R., L'Italien, G., van der Flier, W., Aalten, P., Teunissen, C., Barkhof, F., Blennow, K., Wolz, R., Rueckert, D., Verhey, F., Visser, P.J., 2012. Test sequence of CSF and MRI biomarkers for prediction of AD in subjects with MCI. Neurobiol. Aging 33, 2272–2281.
Wahlund, L.O., Blennow, K., 2003. Cerebrospinal fluid biomarkers for disease stage and intensity in cognitively impaired patients. Neurosci. Lett. 339, 99–102.
Wahlund, L.O., Julin, P., Johansson, S.E., Scheltens, P., 2000. Visual rating and volumetry of the medial temporal lobe on magnetic resonance imaging in dementia: a comparative study. J. Neurol. Neurosurg. Psychiatry 69, 630–635.
Wahlund, L.O., Julin, P., Lindqvist, J., Scheltens, P., 1999. Visual assessment of medial temporal lobe atrophy in demented and healthy control subjects: correlation with volumetry. Psychiatry Res. 90, 193–199.
Wang, H., Su, M.Y., 2006. Regional pattern of increased water diffusivity in hippocampus and corpus callosum in mild cognitive impairment. Dement. Geriatr. Cogn. Disord. 22, 223–229.
Westman, E., Cavallin, L., Muehlboeck, J.S., Zhang, Y., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Spenger, C., Lovestone, S., Simmons, A., Wahlund, L.O., AddNeuroMed consortium, 2011. Sensitivity and specificity of medial temporal lobe visual ratings and multivariate regional MRI classification in Alzheimer's disease. PLoS One 6, e22506.
Wolz, R., Aljabar, P., Hajnal, J.V., Hammers, A., Rueckert, D., 2010a. LEAP: Learning embeddings for atlas propagation. Neuroimage 49, 1316–1325.
Wolz, R., Heckemann, R.A., Aljabar, P., Hajnal, J.V., Hammers, A., Lötjönen, J., Rueckert, D., Alzheimer's Disease Neuroimaging Initiative, 2010b. Measurement of hippocampal atrophy using 4D graph-cut segmentation: application to ADNI. Neuroimage 52, 109–118.
Yavuz, B.B., Ariogul, S., Cankurtaran, M., Oguz, K.K., Halil, M., Dagli, N., Cankurtaran, E.S., 2007. Hippocampal atrophy correlates with the severity of cognitive decline. Int. Psychogeriatr. 19, 767–777.
Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57.