PAPER

Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment

P J Visser, F R J Verhey, P A M Hofman, P Scheltens, J Jolles

J Neurol Neurosurg Psychiatry 2002;**72**:491–497

See end of article for authors' affiliations

Correspondence to: Dr P J Visser, Department of Psychiatry, Academisch Ziekenhuis Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands; pj.visser@np.unimaas.nl

Received 28 February 2000
In final revised form 9 October 2001
Accepted 23 October 2001

Objectives: To investigate whether medial temporal lobe atrophy predicted outcome in patients with minor cognitive impairment and whether assessment of the medial temporal lobe could increase the predictive accuracy of age and delayed recall for outcome. Quantitative and qualitative methods of assessing the medial temporal lobe were also compared.
Methods: Patients with minor cognitive impairment older than 50 years (n=31) were selected from a memory clinic and were followed up for on average 1.9 years. The medial temporal lobe was assessed in three different ways: volumetry of the hippocampus, volumetry of the parahippocampal gyrus, and qualitative rating of medial temporal lobe atrophy (MTA). Outcome measures were Alzheimer type dementia or cognitive decline at follow up. Delayed recall was tested with a verbal learning test.
Results: Ten patients had experienced cognitive decline at follow up, of whom seven had probable Alzheimer type dementia. All medial temporal lobe measurements were associated with cognitive decline at follow up (p trend analysis between 0.001 (hippocampus) and 0.05 (parahippocampal gyrus)). Only the hippocampal volume and MTA score were associated with Alzheimer type dementia at follow up (p trend analysis respectively 0.003 and 0.01). All medial temporal lobe measurements increased the predictive accuracy of age and the delayed recall score for cognitive decline (p increase in predictive accuracy varied between <0.001 (hippocampus) and 0.02 (parahippocampal gyrus and MTA score)) and the hippocampal volume and the MTA score increased the predictive accuracy of age and the delayed recall score for Alzheimer type dementia (p= 0.02).
Conclusions: The ability to detect patients at high risk for Alzheimer type dementia among those with minor cognitive impairment increases when data on age and memory function are combined with measures of medial temporal lobe atrophy. Volumetry of the hippocampus is preferred, but qualitative rating of medial temporal lobe atrophy is a good alternative.

Many patients who are investigated for cognitive impairment are not demented at the time of the examination but some of them may develop Alzheimer type dementia over several years. It is difficult to identify these patients. It is important to select them because they may benefit from drugs that have been shown to improve cognition in patients with probable Alzheimer's disease, or drugs that may slow the progression of the disease.[1] In addition, the caregivers of these patients may benefit from counselling on how to handle the cognitive impairment of their partners. One of the best predictors of Alzheimer type dementia in patients with minor cognitive impairment is memory function,[2–4] but the sensitivity of memory functioning for predicting Alzheimer type dementia was less than 80% in most studies.[2–4] In addition, not all patients with memory impairment develop Alzheimer type dementia and the memory impairment may be reversible.[5] Several studies have indicated that atrophy of the medial temporal lobe is predictive of Alzheimer type dementia in non-demented patients[6–11] and that measures of medial temporal lobe atrophy can improve the predictive accuracy of memory function for Alzheimer type dementia[6 10] or can predict the dementia independently from memory function.[11] Because only two of these studies were performed in a clinical setting, it remains uncertain whether the medial temporal lobe should be evaluated as part of the diagnostic investigation of non-demented patients with mild cognitive impairment. Moreover, as there are different methods of assessing the medial temporal lobe, it is unclear which method has the best predictive accuracy: volumetry of the hippocampus,[9] volumetry of the parahippocampal gyrus,[10] or qualitative assessment of the medial temporal lobe.[6 10]

The aim of the present longitudinal study was to investigate whether medial temporal lobe atrophy predicted outcome in elderly patients with minor cognitive impairment and whether assessment of the medial temporal lobe could increase the predictive accuracy of age and delayed recall for clinical outcome. We compared three different methods of assessing the medial temporal lobe: volumetry of the hippocampus, volumetry of the parahippocampal gyrus, and qualitative scoring of medial temporal lobe atrophy (MTA score). Outcome was defined as Alzheimer type dementia at follow up or cognitive decline at follow up. The second outcome measure not only included patients with Alzheimer type dementia at follow up, but also patients with severe cognitive decline without dementia at follow up.

METHODS

Patients

Patients with minor cognitive impairment were selected from the Maastricht Memory Clinic, a university affiliated outpatient clinic for patients with cognitive impairments.[12] Patients were referred to the clinic by a general practitioner, a neurologist, or a psychiatrist. Inclusion criteria were a score on the global deterioration scale (GDS)[13] of 2 or 3. Exclusion criteria were age below 50 years, a baseline diagnosis of dementia according to the DSM-IV criteria,[14] sensory impairment,

Abbreviations: MTA, medial temporal lobe atrophy; AD, Alzheimer type dementia; GDS, global deterioration scale; HDRS, Hamilton depression rating scale; MMSE, mini mental state examination; AVLT, auditory verbal learning test; SCWT, Stroop colour word test

psychosis, panic disorder, bipolar disorder, a score on the Hamilton depression rating scale-17 items (HDRS)[15] higher than 22,[16] or cognitive problems in relation to cerebrovascular events, neurodegenerative diseases (for example, Parkinson's disease or Huntington's disease), brain neoplasm, head trauma, drug intoxication, alcohol misuse, hypothyroid or hyperthyroid function, or vitamin deficiency. Thirty one patients were included in the study. Some of these patients had vascular risk factors or vascular disorders—that is, hypertension (diastolic blood pressure ⩾95, systolic blood pressure ⩾170 on a single measurement, or treatment for hypertension) (n=8), total cholesterol serum concentrations ⩾6.0 mmol/l (n=5), smoking (n=6), angina pectoris (n=1), transient ischaemic attack (n=2), and lacunar infarction on MRI (n=2). The vascular disorders were thought not to be related to the cognitive impairment because there was no relation between the vascular event and the onset of cognitive impairment, nor was there a sudden onset of cognitive impairment. After the study was explained to them, patients gave their written informed consent.

Baseline assessment and clinical diagnosis

At baseline patients underwent a standardised assessment which included a detailed history provided by the patient and a relevant other person, a psychiatric, neurological, and physical examination, appropriate laboratory tests, and a neuropsychological assessment (see below) as described elsewhere.[12] In addition, the mini mental state examination (MMSE),[17] as a measure of global cognitive impairment, the GDS,[13] which is a scale for staging levels of cognitive impairment, the Blessed dementia rating scale part I,[18] as a measure of functional impairment, and the HDRS,[15] were administered. Psychiatric diagnoses were made according to DSM-IV criteria.[14] The diagnosis of Alzheimer type dementia was made according to the NINCDS-ADRDA criteria.[19] No patient received antidementia drugs.

Follow up assessment

The patients were invited for a follow up assessment between 1 and 3 years after the first assessment. The average follow up period was 1.9 years (SD 0.7). The follow up assessment consisted of a standardised questionnaire about medical history and cognitive complaints, the MMSE, the GDS, the HDRS, and a neuropsychological assessment (see below). The diagnosis at follow up was made by an experienced neuropsychiatrist who was unaware of the results of the baseline assessment including the MRI data. If the patient refused to come for the follow up assessment, a telephone interview was conducted which included a standardised questionnaire about medical history and cognitive complaints (n=1). No neuropsychological testing was done at follow up in seven patients because of refusal (n=5), severe cognitive impairment (n=1), or severe illness (n=1).

The diagnosis of cognitive decline at follow up was made when patients had Alzheimer type dementia or when severe cognitive decline without dementia was present at follow up. Decline in non-demented patients was defined as a negative change of four points or more on the MMSE[20] or decline on the delayed recall task such that both a decline >1 SD on the task was present and the follow up score was below the 10th percentile. The last restriction was taken to exclude patients with regression to the mean. When only the MMSE or the delayed recall score was available at follow up (n=2), the patient was classified according to that score only.

Neuropsychological methodology

The neuropsychological assessment consisted of a series of standard clinical tests,[12 21] including the auditory verbal learning test (AVLT) (a test that assesses immediate and delayed recall of a list of 15 unrelated words),[22 23] the Stroop colour word test (SCWT) (a test that assesses simple (card 1) and complex cognitive speed (card 3)),[24] verbal fluency (the ability to name as many professions/trades as possible within 1 minute), and intelligence.[25] Delayed recall performance of the AVLT was selected because several studies have indicated that this is a strong neuropsychological predictor of Alzheimer type dementia.[2 3] Delayed recall performance was not tested at baseline in two patients because they refused to do the test.

Because the cognitive scores correlated with age, sex, and education, we corrected the scores for these variables. Intelligence was corrected according to published age norms and expressed as an intelligence quotient (IQ).[25] The correction of the other cognitive scores was based on a reference population of 1070 cognitively normal patients older than 50 years who had been randomly selected from a registry of general practitioners as described in detail elsewhere.[26 27] On the basis of the reference population, an expected score for a given age, sex, and level of education was calculated.[5 16] This score was subtracted from the observed score and the residue was divided by the SD of the residue in the reference population to give a z score.[5] The sign of the z scores of the SCWT cards 1 and 3 was inverted such that a z score below zero indicated below average performance. We did not correct the MMSE score, in order to facilitate comparisons with other studies. In addition, the analyses with uncorrected MMSE scores yielded similar results as the analyses with corrected MMSE scores.

MRI methodology

A three dimensional volumetric scan (T1 weighted, fast field echo, TR 24 ms, TE 7 ms, flipangle 30°, number of averages=2, FOV 230 mm, resolution 256×154) and an inversion recovery scan (TR 2107 ms, TE 18 ms, turbofactor=3, flipangle 90°, number of averages=2, FOV 230 mm, resolution 256×177) were performed on a 1.5 Tesla scanner (Gyroscan ACS-II, Philips). The slice thickness of the three dimensional volumetric scan was 1.5 mm and the scan axis was coronal, perpendicular to the intercommissural line. The slice thickness of inversion recovery scan was 3 mm and the scan axis was coronal, perpendicular to the long axis of the hippocampus. The hippocampus, the parahippocampal gyrus, and the intracranial area were measured on the three dimensional volume scan and the MTA score was determined from the inversion recovery scan.

Methodology of brain measurements

Data were transferred to a SUN workstation and the regions of interest were measured with ShowImage (developed at the Department of Clinical Physics and Informatics, VrijeUniversiteit, Amsterdam, The Netherlands).[28] The MR scan of one patient was not available for volumetry and in this patient only the qualitative rating was performed. The brain structures were manually traced with a mouse driven cursor. The volumes of the left side and right side were added. The volume of the brain structure was calculated by multiplying the surface area of each region of interest by the slice thickness and summing the volumes of all slices on which the structure was measured. Measurements were done with reference to an anatomical atlas.[29] The hippocampus and parahippocampal gyrus were measured by one rater and the intracranial area was measured by another rater. All raters were blinded to all clinical information.

Volumetry of the hippocampus, parahippocampal gyrus, and intracranial area

The hippocampus was measured on the slice on which both the semiannular sulcus and a notch between the amygdala and the hippocampus in the medial wall of the lateral ventricle were visible,[29] and then on every second slice. The last slice was that before the slice on which the crura of the fornices were visible. On average 10 slices on each side were measured

(range 8–13). The volume of the parahippocampal gyrus was measured on the same slices, except for the last slice in order not to include the isthmus of the cingulate gyrus. On average nine slices on each side were measured (range 7–12). The intracranial area was measured in a rostrocaudal direction on three slices: on the first slice on which the third ventricle appeared, on the slice on which the mamillary bodies had the largest volume, and on the last slice on which the third ventricle was visible. The anatomical boundaries of the hippocampus, parahippocampal gyrus, and intracranial area have been described in detail elsewhere.[30] Ten scans were remeasured to assess the intraobserver variability. The intraclass correlation coefficient between the first and second measurement was 0.95 for the hippocampus, 0.92 for the parahippocampal gyrus, and 0.99 for the intracranial area.

The volumes of the hippocampus and parahippocampal gyrus were corrected for age, sex, intracranial area, and number of slices.[30] We corrected for the number of slices on which the volume of the hippocampus or parahippocampal gyrus were measured in order to reduce the variance because the number of slices correlated with the total volume of the hippocampus and parahippocampal gyrus but the number of slices did not depend on atrophy of these structures.[30] The correction for age, sex, intracranial area, and number of slices was based on a population of 60 healthy patients aged between 21 and 82 years (average age 56 years (SD 15.9)).[31] Regression analysis was performed with the brain structure as dependent variable and intracranial area, age, number of slices, sex, and the interaction term age by sex (because of reported differences in aging between males and females) as independent variables.[30] Variables and interaction terms that were significant at the p=0.05 level were included in the final model. These variables were age, sex, intracranial area, and number of slices in the model with hippocampal volume as dependent variable and age, intracranial area, and number of slices in the model with parahippocampal gyrus volume as dependent variable. Because age was included in both regression models the data of the reference population could be used for the correction of brain volumes in the study population even though the age range in the reference population was not the same as that in the study population.[31] On the basis of the regression model we calculated z scores in the same way as we did for the delayed recall. We classified the brain volumes on the basis of the z score in tertiles. A z score above 0.44 corresponds to a brain volume in the highest tertile of the reference population, a z score between 0.44 and −0.44 corresponds to a brain volume in the middle tertile of the reference population, a z score below −0.44 corresponds to a brain volume in the lowest tertile of the reference population.

MTA score

The MTA score is based on a visual estimation of the volume of the medial temporal lobe, including the hippocampus proper, dentate gyrus, subiculum, and parahippocampal gyrus, and the volume of the surrounding CSF spaces, in particular the temporal horn of the lateral ventricle and the choroid fissure, on both sides.[32 33] The MTA score ranges from 0 (no atrophy) to 4 (severe atrophy). The MTA was scored by a neurologist who was blinded to all clinical information. The intrarater reliability of this rater was substantial (κ=0.70).[34] The left and right MTA scores were averaged and corrected for age and sex using multiple linear regression on the basis of the same reference population that was used for correcting the hippocampal and parahippocampal gyrus volume. A z score was calculated as described above. The sign of the z scores was inverted such that a z score below zero indicated more than average atrophy. We classified the MTA rating on the basis of the z score in tertiles in the same way as we did for the hippocampus and parahippocampal gyrus.

Statistics

The data were analyzed using SPSS for the Macintosh 4.0 (SPSS Inc, Chigaco, IL, USA). Trend analysis was used to investigate the relation between the tertile scores of hippocampal volume, parahippocampal gyrus volume, or MTA score and change in cognitive function and cognitive outcome. In the trend analyses the tertile scores were considered a continuous variable in a multiple linear regression model. Intraclass correlation coefficients were calculated to assess the intraobserver and interobserver variability.[35] Logistic regression was used to evaluate the predictive value of age (50–60 years, 60–70 years, and >70 years), delayed recall, and the z scores of the hippocampal volume or parahippocampal gyrus volume or MTA score for clinical outcome (Alzheimer type dementia *v* no Alzheimer type dementia, and the absence or presence of cognitive decline at follow up). Age was used as a predictor because it is a risk factor for Alzheimer type dementia that is independent of age corrected brain volumes or age corrected memory scores.[5 11] Age and delayed recall were added in the first two steps and the brain volume or MTA score in the third step. To assess whether the goodness of fit improved after each step, the decrease in deviance (or −2 log likelihood) was tested. All tests were two tailed, and the significance level was set at 0.05.

RESULTS

Baseline characteristics of the study population are listed in table 1. Information on the presence or absence of Alzheimer type dementia was available in 30 patients (97%). Seven patients (22% of the patients with known outcome) were demented and had probable Alzheimer type dementia at follow up. Of the patients with no dementia at follow up (n=23), three had cognitive decline at follow up (two patients with a decline on the MMSE ⩾4, and one patient with a decline on the delayed recall >1 SD), 17 patients had no cognitive decline, and three patients had no cognitive scores at follow up. The patients with Alzheimer type dementia at follow up and the non-demented patients with cognitive decline at follow up will be referred to as the cognitive decline group (n=10).

There was a statistically significant association between the volume of the hippocampus and the MTA score at baseline and

Table 1 Patient characteristics

No	31
Age (y)	64.9 (9.5)
Sex ratio	18M:13F
Education (y)*	10.7 (3.2)
MMSE score	27.7 (1.8)
GDS score:	
2	17
3	14
BDRS score	1.9 (1.9)
HDRS score	9.8 (6.5)
Immediate recall (z score)	0.02 (1.3)
Delayed recall (raw score)	6.1 (3.0)
Delayed recall (z score)	−0.91 (0.84)
SCWT: card 1 (z score)	−1.0 (1.5)
SCWT: card 3 (z score)	−1.84 (1.3)
Verbal fluency (z score)	−0.80 (1.0)
IQ	112.3 (12.5)
Hippocampus (z score)	−0.23 (1.1)
Parahippocampal gyrus (z score)	0.10 (1.1)
MTA score (uncorrected, average of left and rigth side)	1.1 (1.1)
MTA score (z score)	−0.35 (1.4)

Values are means (SD).
*Years spent in primary, secondary and higher level education.
MMSE, mini mental state examination; GDS, global deterioration scale; HDRS, Hamilton depression rating scale; BDRS, Blessed dementia rating scale; SCWT, Stroop colour word test, IQ, intelligence quotient; MTA, medial temporal lobe atrophy.

Table 2 Baseline and follow up data for parahippocampal volume, hippocampal gyrus volume, and MTA score

	Hippocampus volume at baseline*			p Value trend analysis
	Highest tertile	Middle tertile	Lowest tertile	
(A) Baseline and follow up data according to hippocampal volume at baseline:				
Number	9	9	12	
Baseline data				
Age (y)	60.6 (8.6)	65.9 (8.4)	66.9 (10.9)	0.15
MMSE score	27.8 (1.7)	27.9 (1.7)	27.3 (2.1)	0.50
Delayed recall (z score)	–1.15 (0.84)	–0.42 (0.84)†	–1.1 (0.80)†	0.94
Follow up data				
Change MMSE	0.0 (2.2)§	–0.78 (1.7)	–4.7 (4.4)§	0.002
Change delayed recall	0.73 (1.2)§	–0.33 (1.2)†	–0.29 (1.5)¶	0.18
AD/no AD at follow up (% AD)	0/9 (0)	1/8 (13)	6/5 (55)†	0.003
CD/no CD at follow up (% CD)	0/6 (0)§	2/7 (30)	8/3 (72)†	0.001
	Parahippocampal gyrus volume at baseline*			**p Value trend analysis**
	Highest tertile	**Middle tertile**	**Lowest tertile**	
(B) Baseline and follow up data according to parahippocampal gyrus volume at baseline:				
Number	12	10	8	
Baseline data				
Age (y)	64.2 (10.4)	64.2 (9.3)	66.1 (10.0)	0.69
MMSE score	27.0 (2.0)	28.2 (1.9)	27.8 (1.7)	0.31
Delayed recall (z score)	–0.93 (0.92)‡	–0.80 (0.93)	–1.05 (0.76)	0.80
Follow up data				
Change MMSE	–0.86 (2.3)**	–1.1 (3.1)	–4.6 (4.5)†	0.05
Change delayed recall	0.61 (1.3)**	–0.11 (1.2)†	–0.64 (1.4)‡	0.09
AD/no AD at follow up (% AD)	2/10 (17)	2/8 (20)	3/4 (43)†	0.24
CD/no CD at follow up (% CD)	2/7 (22)§	3/7 (30)	5/2 (71)†	0.05
	MTA score at baseline*			**p Value trend analysis**
	Highest tertile	**Middle tertile**	**Lowest tertile**	
(C) Baseline and follow up data according to MTA score at baseline:				
Number	12	6	13	
Baseline data				
Age (y)	59.2 (8.1)	67.9 (2.0)	68.8 (10.1)	0.01
MMSE score	28.3 (1.4)	27.8 (1.9)	27.0 (2.0)	0.08
Delayed recall (z score)	–0.86 (0.72)†	–0.58 (1.1)	–1.1 (0.80)†	0.46
Follow up data				
Change MMSE score	–0.73 (2.1)†	–1.2 (1.6)†	–3.9 (5.0)¶	0.05
Change delayed recall score	0.11 (1.4)§	0.14 (1.4)†	–0.16 (1.3)¶	0.67
AD/no AD at follow up (% AD)	1/11 (8)	0/6 (0)	6/6 (50)†	0.01
CD/no CD at follow up (% CD)	2/9 (18)†	0/5 (0)†	8/3 (73)‡	0.01

Values are means (SD). *Tertiles are based on a reference population of healthy subjects. A low tertile score indicates a small hippocampal or parahippocampal gyrus volume, or an high MTA score; †one patient with missing data; ‡two patients with missing data; §three patients with missing data; ¶four patients with missing data; **five patients with missing data.
MMSE, Mini mental state examination; AD, Alzheimer type dementia; CD, cognitive decline: patients with Alzheimer type dementia at follow up or patients with severe cognitive decline at follow up without dementia.

decline on the MMSE, the diagnosis of Alzheimer type dementia, and the diagnosis of cognitive decline at follow up. The volume of the parahippocampal gyrus was significantly associated with decline on the MMSE and cognitive decline at follow up, and tended to be associated with change in delayed recall performance at follow up (table 2). These associations remained similar after adjustments for duration of follow up. The association between medial temporal lobe atrophy and change of cognitive scores remained the same after correction for the baseline cognitive scores. None of the measures of medial temporal lobe atrophy was associated with age or cognitive scores at baseline except for the MTA score that was associated with age at baseline.

In the logistic regression analyses three patients were excluded because of missing data: two patients had no delayed recall score at baseline and in one patient medial temporal lobe atrophy was only assessed with the qualitative rating scale and not with volumetry. The first set of logistic regression analyses was performed with Alzheimer type dementia at follow up as dependent variable and included data for six patients with Alzheimer type dementia at follow up and 21 non-demented patients. After age and the delayed recall score were entered in the first two steps, the hippocampus volume and the MTA score improved the model (table 3). In the second set of logistic regression analyses, the dependent variable was cognitive decline at follow up and included data for nine patients with cognitive decline at follow up and 15 patients without cognitive decline. After age and the delayed recall score were entered in the first two steps, the volumes of the hippocampus and the parahippocampal gyrus, and the MTA score all improved the model (table 3). The decrease in deviance was largest for the hippocampal volume, indicating that this variable increased the predictive accuracy the most.

DISCUSSION

The main findings of this prospective study of patients with minor cognitive impairment in a clinical setting were that measures of medial temporal atrophy predicted outcome at follow up and improved the predictive accuracy of age and delayed recall performance for outcome. The hippocampal volume was a better predictor of outcome than the MTA score, and the MTA score was a better predictor than the parahippocampal gyrus volume.

Table 3 Logistic regression analyses of clinical outcome

	Change in deviance*	p Value change deviance	Overall accuracy (%)†	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
AD as outcome:							
Step 1 Age	3.2	0.07	78	0	100	–	78
Step 2 Delayed recall	3.2	0.07	74	17	90	33	79
Step 3 HC	5.1	0.02	81	50	90	60	86
Step 3 PHG	1.4	0.25	74	33	86	40	82
Step 3 MTA score	5.5	0.02	81	50	90	60	86
Cognitive decline as outcome:							
Step 1 Age	4.6	0.03	71	44	87	67	72
Step 2 Delayed recall	2.6	0.11	71	73	67	60	79
Step 3 HC	24.6	<0.001	100	100	100	100	100
Step 3 PHG	5.2	0.02	75	56	87	71	76
Step 3 MTA score	5.4	0.02	83	78	87	78	87

*Change in deviance is the change in deviance from maximum deviance (step 1) or the previous step (steps 2 and 3); †the overall accuracy, sensitivity, specificity, PPV and NPV were calculated on the basis of the predicted probability of outcome: Alzheimer type dementia or cognitive decline were considered to be predicted by the model if the predicted probability of Alzheimer type dementia or cognitive decline was ⩾0.5.
HC, hippocampus; PHG, parahippocampal gyrus; MTA, medial temporal lobe atrophy; PPV, positive predictive value; NPV, negative predictive value; –,could not be calculated because no patients were predicted to have Alzheimer type dementia at follow up.

The finding that medial temporal lobe atrophy was associated with cognitive outcome is consistent with earlier observations of patients with or without mild cognitive impairment.[6–11] All medial temporal lobe measures were found to improve the predictive accuracy of age and the delayed recall score for cognitive decline at follow up. These findings corroborate the finding that assessment of the medial temporal lobe increases predictive accuracy for clinical outcome above that of cognitive dysfunction[6 10] and are consistent with the finding that hippocampal volume and memory scores were independent predictors of Alzheimer type dementia in patients with minor cognitive impairment.[11]

The trend analyses and logistic regression analyses indicated that the hippocampal volume was a better predictor of outcome than the volume of the parahippocampal gyrus. This corroborates the findings of Kaye *et al*, who showed that non-demented patients with Alzheimer type dementia at follow up had at baseline smaller hippocampal volumes but not parahippocampal gyrus volumes at baseline than patients without dementia at follow up,[9] but it is by contrast with our previous study that showed a better predictive accuracy of the parahippocampal gyrus for Alzheimer type dementia in non-demented elderly people.[10] This discrepancy may have resulted from the difference in criteria for minor cognitive impairment that have been used. In our previous study, patients were selected according to the criteria of minimal dementia[10] and these patients had at baseline more severe cognitive impairment (average MMSE score of 22.6) than the patients with mild cognitive impairments in the present study (average MMSE score 27.7) or the study of Kaye *et al* (average MMSE score 26.9).[9] Another explanation for the lower discriminative ability of the parahippocampal gyrus may be the larger interindividual differences in this structure.[36] The fact that in our previous study the parahippocampal gyrus volume was nevertheless a better predictor than the hippocampal volume, was perhaps because measurements were performed on only four slices. This may have decreased the accuracy with which the hippocampus was measured because the shape of the hippocampus changes along the longitudinal axis. By contrast, the parahippocampal gyrus has a more stable form. It is therefore possible that the measurement error of the hippocampus outweighed the interindividual variability of the parahippocampal gyrus. The hippocampal volume was a better predictor of Alzheimer type dementia and cognitive decline in the trend analyses than the MTA score. In the multivariate analyses, the hippocampal volume and the MTA score could predict Alzheimer type dementia with the same accuracy but the hippocampal volume was a better predictor of cognitive decline than the MTA score. These findings suggest that volumetry of the hippocampus should be preferred above the MTA score in predicting cognitive outcome. However, volumetry of the hippocampus takes about 15 minutes and data for a reference population of normal patients are needed to correct for differences in intracranial volume. This will probably limit the use of volumetry in a clinical setting. The MTA score has the advantage that it can be scored on hard copies and that it takes only a few minutes to perform. Because the MTA score has a good predictive accuracy, it could be used in settings where volumetry is not possible. One possible disadvantage of the MTA score is the interrater variability. In a previous study it was shown that the interrater agreement on MTA ratings of 100 scans of four raters was fair to substantial (κ between 0.34 and 0.57), which indicates that the generalisation of the MTA score may be less satisfactory.[34] To prevent interrater variability in MTA scoring raters should therefore be well trained.

The differences in predictive accuracy between the three measures of medial temporal lobe atrophy suggest that these measures assess different aspects of medial temporal lobe atrophy. To investigate this possibility we correlated in the present sample post hoc the z scores of the three medial temporal lobe measures with each other. The correlation coefficient between the volumes of the hippocampus and the parahippocampal gyrus was 0.69 ($p<0.001$), the correlation coefficient between the volume of the hippocampus and the MTA score was 0.29 ($p=0.11$), and the correlation coefficient between the volume of the parahippocampal gyrus and the MTA score was 0.26 ($p=0.17$). Because the patients in these correlations were their own controls, we repeated the correlations using brain volumes and MTA scores that were not corrected for age. Without age correction, the correlation coefficient between the volumes of the hippocampus and the parahippocampal gyrus remained the same but the correlation coefficients between the volumetric measures and the MTA score were higher ($r=0.42$, $p=0.02$ for both the correlation between the volume of the hippocampus and the MTA score and the correlation between the volume of the parahippocampal gyrus and the MTA score). The correlation coefficient between the volume of the hippocampus and the MTA score was similar to that reported in a previous study.[37] These data indicate that the three measures of medial temporal lobe atrophy indeed assess different aspects of medial temporal lobe atrophy. One possible explanation of the moderate correlation between the MTA score and the volumetric measures is that

the MTA score is based not only on a visual estimation of the volume of the medial temporal lobe, but also of the volume of the surrounding CSF spaces, in particular the temporal horn of the lateral ventricle and the choroid fissure. Evidence for this explanation comes from the finding that the MTA score correlated significantly with the volume of the temporal horn of the lateral ventricle in the present sample (r=0.51, p=0.003, if both measures were corrected for age, and r=0.69, p<0.001, if both measures were not corrected for age).

Minor cognitive impairment in the present study was defined as cognitive impairment that led to a referral to a memory clinic but that was not severe enough to meet the criteria of dementia which equals GDS stages 2 and 3. There are also other concepts of minor cognitive impairment, such as mild cognitive impairment (MCI).[38] In the present study, 19 patients met the criteria of MCI (two patients did not meet the criteria because cognitive test data were missing and 10 patients because they were younger than 60 years). Trend analyses and logistic regression analyses with these patients yielded similar results as the analyses with the whole sample which indicates that the present findings will also apply in patients with MCI. The mean age in the study sample (65 years) was lower than that of other studies investigating the predictive accuracy of medial temporal lobe atrophy (mean age range from 71 to 78).[6 10 11] This may have negatively influenced the positive predictive value of medial temporal lobe atrophy because the conversion rate to dementia is lower in younger patients than it is in older patients. We corrected for this difference in conversion rate by including age in the multivariate models. We also performed analyses with a subgroup of patients older than 60 years (average 71 years) (data not shown) which yielded similar results as those with the whole sample. We did not exclude patients with mild to moderate depression because mild to moderate depression is often seen in patients with preclinical Alzheimer type dementia.[16] Because depression may cause cognitive impairment we may have misclassified depressed patients as having cognitive decline at follow up. This seems unlikely because patients with cognitive decline at follow up who had high depression scores at baseline (HDRS score of 19 and 21) had improvement of the depression scores at follow up (average improvement of five points). In addition, trend and logistic regression analyses with correction for depression severity yielded similar results as the analyses without correction. We included patients with cardiovascular risk factors or vascular disorders that might have contributed to cognitive impairment or to conversion to dementia. However, adjustment for these risk factors or vascular disorders had no effect on the results (data not shown). Also exclusion of the patients with a history of a transient ischaemic attack or a lacunar infarction yielded similar results.

One of the limitations of the study was the short follow up period. We may therefore have missed patients who would have become demented after the follow up assessment. For this reason we also used a broader definition of cognitive decline as an outcome measure, but it remains to be investigated whether the patients with cognitive decline at follow up who were not demented have since developed Alzheimer type dementia. Pathological confirmation of the clinical diagnosis at follow up was not possible and the pathological diagnosis remains therefore uncertain. The small sample size may have limited the ability to detect significant differences. The hippocampal volume and parahippocampal gyrus volumetry and MTA scoring were measured on MRI with a different scan axis. Because the slice thickness was thin it seems unlikely that the difference in scan axis has introduced major bias.We have investigated whether assessment of the medial temporal lobe could increase the predictive accuracy of age and delayed recall for clinical outcome but it remains to be investigated which will be the simplest model for predicting outcome.

The clinical relevance of these findings is that the ability to detect patients with minor cognitive impairment who are at high risk for Alzheimer type dementia will increase when data on age and memory function are combined with measures of medial temporal lobe atrophy. Assessment of medial temporal lobe atrophy may therefore be a useful supplement to the diagnostic investigation of patients with minor cognitive impairment.

ACKNOWLEDGEMENTS

We thank D Tisserand for assisting in the measurement of the brain structures, Dr A Kester for statistical advice, and Dr J Sykes for linguistic advice.

Authors' affiliations

P J Visser, F R J Verhey, J Jolles, Institute of Brain and Behavior, Department of Psychiatry, University of Maastricht, Maastricht, The Netherlands
P A M Hofman, Department of Radiology, Academisch Ziekenhuis Maastricht, Maastricht, The Netherlands
P Scheltens, Department of Neurology, VU Medical Centre, Amsterdam, The Netherlands

REFERENCES

1 **Felician O**, Sandson T. The neurobiology and pharmacotherapy of Alzheimer's disease. *J Neuropsychiatry Clin Neurosci* 1999;**11**:19–31.
2 **Masur DM**, Sliwinski M, Lipton RB, *et al*. Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology* 1994;**44**:1427–32.
3 **Tierney MC**, Szalai JP, Snow WG, *et al*. Prediction of probable Alzheimer's disease in memory impaired patients. *Neurology* 1996;**46**:661–5.
4 **Sliwinski M**, Buschke H, Stewart W, *et al*. The effect of dementia risk factors on comparative and diagnostic selective reminding norms. *J Int Neuropsychol Soc* 1997;**3**:317–26.
5 **Visser PJ**, Verhey FRJ, Ponds RWHM, *et al*. Course of objective memory impairment in non-demented subjects attending a memory clinic and predictors of outcome. *Int J Geriatr Psychiatry* 2000;**15**:363–72.
6 **de Leon MJ**, George AE, Convit A, *et al*. The radiologic prediction of Alzheimer disease: the atrophic hippocampal formation. *AJNR Am J Neuroradiol* 1993;**14**:897–906.
7 **Golomb J**, Kluger A, de Leon M, *et al*. Hippocampal formation size predicts declining memory performance in normal aging. *Neurology* 1996;**47**:810–3.
8 **Fox N**, Warrington E, Freeborough P, *et al*. Presymptomatic hippocampal atrophy in Alzheimer's disease. *Brain* 1996;**119**:2001–7.
9 **Kaye J**, Swihart T, Howieson D, *et al*. Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia. *Neurology* 1997;**48**:1297–304.
10 **Visser PJ**, Scheltens P, Verhey FRJ, *et al*. Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment. *J Neurol* 1999;**246**:477–85.
11 **Jack CR**, Petersen RC, Xu YC, *et al*. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology* 1999;**52**:1397–403.
12 **Verhey FRJ**, Jolles J, Ponds RWHM, *et al*. Diagnosing dementia: a comparison between a monodisciplinary and multidisciplinary approach. *J Neuropsychiatry Clin Neurosci* 1993;**5**:78–85.
13 **Reisberg B**, Ferris SH, De Leon MJ, *et al*. The global deterioration scale for assessment of primary degenerative dementia. *Am J Psychiatry* 1982;**139**:1136–9.
14 **American Psychiatric Association A**. *Diagnostic and statistical manual of mental disorders*. IVth ed. Washington, DC: APA, 1994.
15 **Hamilton M**. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;**23**:56–62.
16 **Visser PJ**, Verhey FRJ, Ponds RWHM, *et al*. Distinction between preclinical dementia and depression. *J Am Geriatr Soc* 2000;**48**:479–84.
17 **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.
18 **Blessed G**, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile changes in the cerebral grey matter of elderly subjects. *Br J Psychiatry* 1968;**114**:797–811.
19 **McKhann G**, Drachman D, Folstein M, *et al*. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work-Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology* 1984;**34**:939–44.
20 **Schmand B**, Lindeboom J, Launer L, *et al*. What is a significant score change on the mini-mental state examination? *Int J Geriatr Psychiatry* 1995;**10**:411–4.
21 **Jolles J**. The early diagnosis of dementia: a possible contribution of neuropsychology. In: Gispen W, Traber J, eds. *Aging of the brain*. Berlin: Springer Verlag; 1986:84–100.

22 **Lezak M**. *Neuropsychological assessment.* 3rd ed. New York: Oxford University Press, 1995.
23 **Brand N**, Jolles J. Learning and retrieval rates of words presented auditorily and visually. *J Gen Psychol* 1985;**112**:201–10.
24 **Stroop J**. Studies of interference in serial verbal reactions. *J Exp Psychol* 1935;**18**:643–62.
25 **Luteyn F**, van der Ploeg FAE. *Groninger intelligentie test.* Lisse, The Netherlands: Swets en Zeitlinger, 1983.
26 **Jolles J**, Houx P, van Boxtel MPJ, *et al*, eds. *Maastricht aging study: determinants of cognitive aging.* Maastricht: Neuropsych Publishers, 1995.
27 **van Boxtel MPJ**, Buntinx F, Houx PJ, *et al*. The relation between morbidity and cognitive performance in a normal aging population. *J Gerontol Med Sci* 1998;**53A**:M147–54.
28 **van Walderveen MMA**, Barkhof F, Hommes OR, *et al*. Correlating MRI and clinical disease activity in multiple sclerosis: relevance of hypointense lesions on short-TR/short-TE (T1-weighted) spin-echo images. *Neurology* 1995;**45**:1684–90.
29 **Duvernoy H**. *The human hippocampus. An atlas of applied anatomy.* München: JF Bergmann Verlag, 1988.
30 **Visser PJ**, Krabbendam L, Verhey FRJ, *et al*. Brain correlates of memory dysfunction in alcoholic Korsakoff's syndrome. *J Neurol Neurosurg Psychiatry* 1999;**67**:774–8.
31 **Tisserand DJ**, Visser PJ, van Boxtel MP, *et al*. The relation between global and limbic brain volumes on MRI and cognitive performance in healthy individuals across the age range. *Neurobiol Aging* 2000;**21**:569–76.
32 **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.
33 **Barber R**, Gholkar A, Scheltens P, *et al*. Medial temporal lobe atrophy on MRI in dementia with Lewy bodies. *Neurology* 1999;**52**:1153–8.
34 **Scheltens P**, Launer L, Barkhof F, *et al*. Inter-observer reliability of visual assessment of hippocampal atrophy on MRI. *J Neurol* 1995;**242**:557–60.
35 **Shrout PE**, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;**86**:420–8.
36 **Jack C**, Petersen R, Xu Y, *et al*. Medial temporal lobe atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology* 1997;**49**:786–94.
37 **Wahlund LO**, Julin P, Lindqvist J, *et al*. Visual assessment of medial temporal lobe atrophy in demented and healthy control subjects: correlation with volumetry. *Psychiatry Res* 1999;**90**:193–9.
38 **Petersen R**, Smith G, Waring S, *et al*. Mild cognitive impairment. Clinical characterization and outcome. *Arch Neurol* 1999;**56**:303–8.