Linear measures of atrophy in mild Alzheimer disease.

G B Frisoni, A Beltramello, C Weiss, C Geroldi, A Bianchetti and M Trabucchi

AJNR Am J Neuroradiol 1996, 17 (5) 913-923
http://www.ajnr.org/content/17/5/913

Linear Measures of Atrophy in Mild Alzheimer Disease

Giovanni B. Frisoni, Alberto Beltramello, Claudia Weiss, Cristina Geroldi, Angelo Bianchetti, and Marco Trabucchi

PURPOSE: To assess the sensitivity of linear measures of brain atrophy in the diagnosis of Alzheimer disease (AD) in the early stages. **METHODS:** Linear measures of regional frontal (bifrontal index, interhemispheric fissure width), medial temporal lobe (interuncal distance, minimum thickness of the medial temporal lobe), and hippocampal (hippocampal height, width of the choroid fissure, width of the temporal horn) atrophy were made on magnified MR images obtained in 46 patients with AD (33 with mild severity and 13 with moderate severity) and in 31 control subjects. Gaussian modeling was used to compute sensitivity with specificity set at 95%. Discriminant analysis was used to identify measures independently contributing to the ability to discriminate AD patients from control subjects. **RESULTS:** The measure with the best sensitivity in discriminating AD patients from control subjects was the width of the temporal horn. A compound measure of width of the temporal horn, width of the choroid fissure, height of the hippocampus, and interuncal distance could discriminate patients with mild AD from control subjects with 86% sensitivity. Cross validation in patients with moderate AD confirmed the usefulness of the model (81% sensitivity). Measures of hippocampal atrophy alone could discriminate patients with mild AD from control subjects with 83% sensitivity; in patients with moderate AD, cross validation produced 87% sensitivity. **CONCLUSIONS:** Linear measures of hippocampal atrophy can be a useful adjunct in the routine diagnosis of AD, even in its early stages.

Index terms: Brain, atrophy; Brain, measurements; Dementia

AJNR Am J Neuroradiol 17:913–923, May 1996

Diagnosis in life of Alzheimer disease (AD) is made on clinical grounds (1), and currently used criteria are burdened with considerable subjective judgments (2), which carry an overall accuracy of 81% to 88% (3). Given the high prevalence of the disease and the increasing treatment options (4), simple and sensitive quantitative indicators of the disease in its early stages might represent an important clinical tool.

In recent years, the quantitative evaluation of cerebral tissue was considered unreliable for the detection of AD (5). However, attention was directed toward indicators of global atrophy, while recent research (6–19) has suggested that the measurement of regional atrophy in the structures of the medial temporal lobe (hippocampus, hippocampal formation, perihippocampal structures) is a promising way to discriminate AD patients from healthy control subjects. Available studies have measured regional atrophy in moderately to severely demented patients (9), or have used accurate but demanding imaging analysis techniques (8, 17, 20) or subjective ratings of atrophy (10, 11). On the other hand, linear measures of atrophy of the medial temporal lobe are both objective and simple enough to be used in routine clinical applications. In line with previous data showing that linear measures of hippocampal atrophy yield satisfactory sensitivity in the detection of AD in the moderate stage (15, 19), in this study we assessed their usefulness in the detection of AD in the mild stages.

Received June 13, 1995; accepted after revision November 15.

From the Alzheimer's Disease Unit, Ospedale S. Cuore FBF, and the Geriatric Research Group, Brescia, Italy (G.B.F., C.G., A.Bi., M.T.); and the Institute of Radiology, University of Verona (Italy) (A.Be., C.W.).

Address reprint requests to Giovanni B. Frisoni, MD, Alzheimer's Disease Unit, Ospedale S. Cuore FBF, via Pilastroni 4, 25123 Brescia, Italy.

AJNR 17:913–923, May 1996 0195-6108/96/1705–0913

Subjects and Methods

Subjects

Our study group included 46 patients with AD (33 with mild severity and 13 with moderate severity) and 31 healthy control subjects. All patients and control subjects

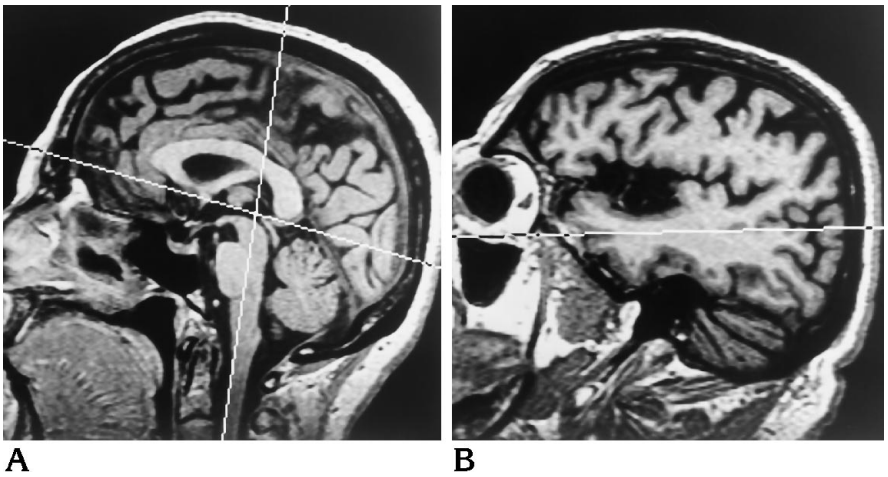

Fig 1. Sagittal three-dimensional gradient-echo MR images (10/4/2, 10° flip angle) of a patient with Alzheimer disease. Midsagittal image (*A*) shows the bicommissural plane and the brain stem axis plane; parasagittal image (*B*) shows the temporal lobe plane.

were recruited at the Alzheimer's Disease Unit, Brescia, Italy, from September 1, 1993, to December 15, 1994. None of the patients or the control subjects was under steroid therapy or had a history of alcohol abuse.

All patients met National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria for probable AD (1). Patients meeting these criteria but with clinical features suggesting non-Alzheimer forms of degenerative dementia (frontal lobe degeneration of non-Alzheimer type, dementia of the Lewy body type, Pick disease) (21–24) were not included in the study.

Subjects with mild and moderate AD were culled strictly from consecutively seen patients whose scores on the Mini Mental State Examination (MMSE) (25) were 20 or greater (mild disease) and between 12 and 19 (moderate disease). Patients with mild AD scored between 0.5 and 1, and those with moderate AD scored 2 on a scale grading overall severity of dementia (clinical dementia rating) (26), which is consistent with severity as defined by means of the MMSE. A complete history, including assessment of basic and instrumental activities of daily living (27, 28), was taken from a proxy informant. Laboratory studies included complete blood count, chemistry profile, chest X-ray, thyroid function, measurements of B_{12} and folic acid, electrocardiogram, electroencephalogram, and computed tomography. Neurologic examination was performed by a neurologist, and physical examination of all systems by a geriatrician. Neuropsychological testing (29) was performed by a psychologist and included MMSE, verbal (logical memory test) and nonverbal (recall of Rey-Osterreith and Wechsler Memory Scale figures) learning, immediate memory (digit span forward and backward), abstract thinking (Raven's progressive matrices), visuospatial planning (clock drawing test), constructional apraxia (copy of Rey-Osterreith and Wechsler Memory Scale figures), verbal fluency (letter and category word fluency) and comprehension (token test), and mood (Geriatric Depression Scale).

The 31 control subjects consisted of relatives of the patients (mostly spouses) who had no detectable cognitive deficit and no history of neurologic disease, although some reported mild subjective memory problems that did not result in impairment in daily activities. All were given the MMSE and were judged not demented by a neurologist and a psychologist involved in the examination of the patients. Apolipoprotein E phenotyping was performed with isoelectric focusing on delipidated plasma samples (30).

Written informed consent was obtained from the control subjects and from the patients or their primary caregivers, after discussion of risks and benefits of the participation. No compensation was provided.

Magnetic Resonance (MR) Imaging Technique and Analysis

MR imaging was performed in the radiology department of the University of Verona with a 1.5-T unit and a standard head coil. A gradient-echo three-dimensional technique was used for image acquisition with parameters of 10/4/2 (repetition time/echo time/excitations), 10° flip angle, 250-mm field of view, 160 × 256 matrix, and inversion time of 300, allowing reconstruction of 1.3-mm-thick contiguous sections. Total acquisition time was 7 minutes 40 seconds. All linear measurements were obtained by the same neuroradiologist on magnified images (magnification factor, 1.5 to 1.7) with the built-in distance measurement software, who was blinded to the diagnosis, age, and sex of the subject.

The following planes were identified (Fig 1): the bicommissural plane, on the midsagittal section, joining the anterior with the posterior commissure (although the anterior commissure is a precise anatomic landmark, the posterior commissure was set at the level of the cranial extremity of the superior collicula) (Fig 1A); the brain stem axis plane, on the midsagittal section, parallel to the dorsal surface of the brain stem (Fig 1A); and the temporal lobe plane, on the parasagittal section where the temporal lobe was best

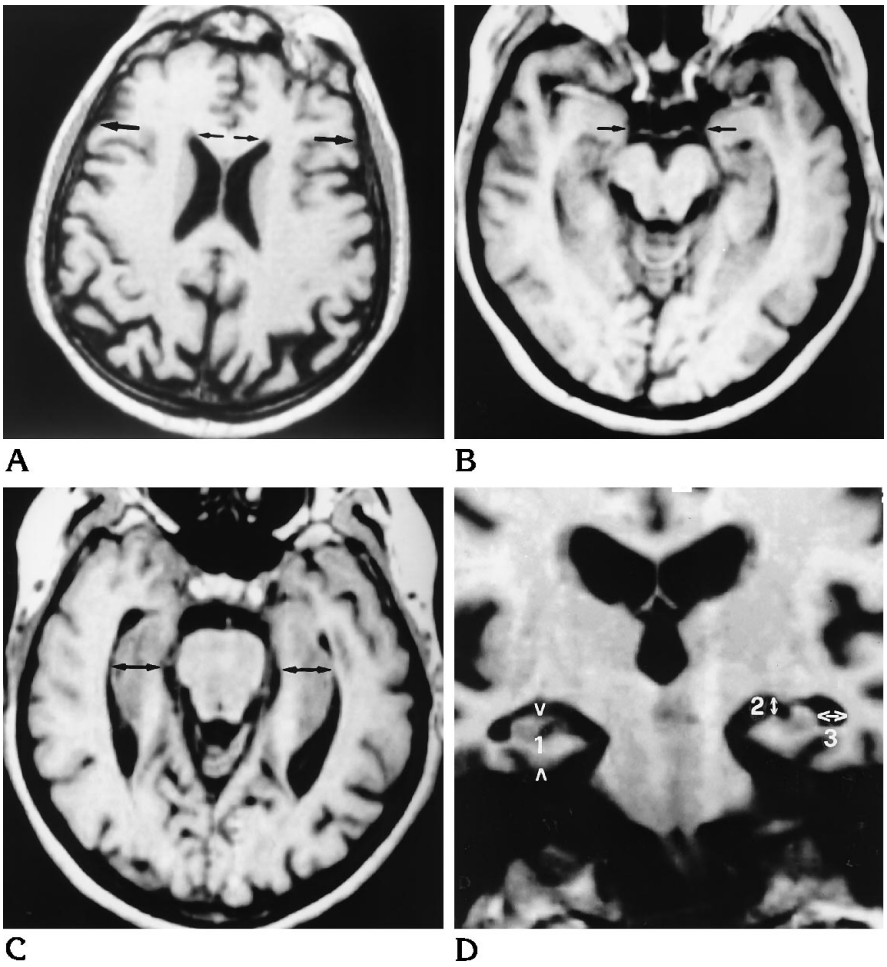

Fig 2. Axial (*A–C*) and coronal (*D*) three-dimensional gradient-echo MR images (10/4/2, 10° flip angle) of a patient with Alzheimer disease. The width between the frontal horns of the lateral ventricles (*smaller arrows*) and the cranial width (*larger arrows*) (bifrontal index) are shown in *A*; *arrows* in *B* indicate interuncal distance; *arrows* in *C* show minimum thickness of the medial temporal lobe; *D* shows hippocampal height (*1*), width of the choroid fissure (*2*), and width of the temporal horn (*3*).

seen in its full length, approximately 20° caudad to the orbitomeatal line (Fig 1B).

The linear measurements that were taken, on both sides when appropriate (Fig 2), are as follows. The bifrontal index, measured on a plane parallel to the temporal lobe plane at the level of the maximal width between the tips of the frontal horns of the lateral ventricles (31), and defined as the ratio of this measure to the width of the brain at the same level multiplied by 100 (Fig 2A) (31). The interhemispheric fissure width, measured on the same plane as the bifrontal index, and defined as the greatest distance between the mesial aspects of the cerebral cortex in the interhemispheric fissure. The interuncal distance, measured on a plane parallel to the bicommissural plane at the level of the suprasellar cistern, as the distance between the unci of the temporal lobes (32) (Fig 2B). The minimum thickness of the medial temporal lobe, measured on a plane parallel to the temporal lobe plane, as the thickness of the medial temporal lobe considered at its narrowest point on the scan section that best represents the medial temporal lobe between its superior and inferior limits (further details can be found in the original article by Jobst et al [9]). The hippocampal height, measured on a plane parallel to the brain stem axis plane where the hippocampal formation was highest, and defined as the greatest height of the hippocampal formation (dentate gyrus, hippocampus proper, and subiculum, together with the parahippocampal gyrus) (10) (Fig 2D). Cranial width (the maximum distance between the inner tables of the calvaria) was measured on the same plane used for hippocampal height measurement. The width of the choroid fissure, measured on the same plane used for hippocampal height measurement, and defined as the vertical width of the choroid fissure centered on the midpoint of the hippocampal formation (10) (Fig 2D). This point usually lies on the line where hippocampal height is taken. And, finally, the width of the temporal horn, measured on the same plane used for hippocampal height measurement (10) (Fig 2D).

For bilateral measures, only the right or left value indicating greater atrophy was considered in the statistical analysis described below.

Test-retest reliability of all measures was assessed in randomly selected patients (n = 10) and control subjects (n = 10) by the same neuroradiologist, who obtained measurements on two separate occasions 2 to 6 weeks apart and who was blind to previous results. Intraclass

correlation coefficient (33) ranged from 0.91 to 0.98 for all measures, indicating good reliability.

Statistical Analysis

Statistical analysis was performed with the Statistical Package for the Social Sciences (34). Differences of continuous or dichotomous clinical, demographic, and radiologic variables between patients and control subjects were assessed with Student's t test or the χ^2 test where appropriate. Normality of the distribution of radiologic measures was assessed with Lilliefors's modification of the Komolgorov-Smirnov test (34). The relationship of measures of brain atrophy with age and cranial width was assessed with Pearson's correlation coefficient.

The increase of atrophy seen in control subjects through mild and moderate AD patients was assessed with an age-adjusted and sex-adjusted test for trend in multiple regression models in which measures of atrophy were dependent and group status (coded as an ordinal variable: 0 = control, 1 = mild AD, and 2 = moderate AD), age, and sex were the independent variables. The measure of association was nonstandardized β with a 95% confidence interval (CI), indicating the adjusted average change of the variable through each group. Significance was that of the associated t statistics.

The effect of sex in the relationship between atrophy and age in the control subjects was addressed with general factorial analysis of variance (ANOVA) models. General factorial ANOVA models were built with sex and age and their interaction as factors.

The normal effect of age (and cranial width when the case) on brain measures was taken into account by transforming measures into multiples of the median (MoM), defined as the ratio of the observed measure to the expected value (9). This last value was computed by regressing brain measures on age and, when appropriate, cranial width in the control subjects. Figure 3 graphically shows the meaning of the MoM of the minimum thickness of the smallest medial temporal lobe. The value of MoM and the relative expected sensitivity best discriminating AD patients from control subjects with 95% specificity were then computed by fitting MoM values to gaussian models in AD patients and in control subjects. A graphic example is provided in Figure 4.

Measures of atrophy independently contributing to the prediction of disease status (AD or control subject) were identified with multivariable discriminant analysis (34) with stepwise selection of variables. This technique minimizes the overlapping between two groups (for example, patients with mild AD and control subjects) by computing a multivariable function that allows computation of a (discriminant) score for each subject. The discriminant scores are such that their mean is highest in one group and lowest in the other, with the smallest possible overlapping and resulting in approximately maximal sensitivity and specificity. Multivariable discriminant analysis takes into account the independent contribution of each variable to the separation of the two groups, and variables unable to increase separation of the two groups are excluded from the final model. Entering of variables in the discriminant model was based on the smallest λ of the discriminant function and on F-to-enter for λ greater than 3.84. Re-

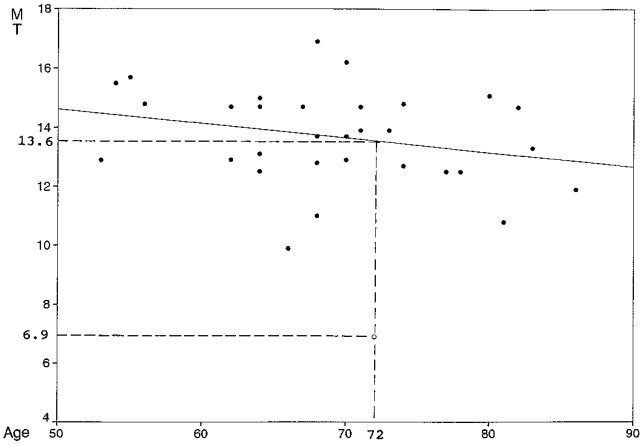

Fig 3. The multiple of the median (MoM) of the minimum thickness of the smallest medial temporal lobe of a patient with mild Alzheimer disease (AD). *MT* indicates minimum thickness of the smallest medial temporal lobe in millimeters; *open circle*, patient with mild AD; *solid circles*, control subjects. The patient is 72 years old, and his MT is 6.9 mm. The expected normal value of MT in a 72-year-old control subject is 13.6 mm. The value of the MoM is 6.9/13.6 = 0.51. Direct measures of atrophy, such as MT and hippocampal height, indicate greater atrophy than that of control subjects for MoM values lower than 1, and indirect measures, such as bifrontal index, width of interhemispheric fissure, interuncal distance, width of the choroid fissure, and width of the temporal horn, indicate greater atrophy than that of control subjects for MoM values greater than 1.

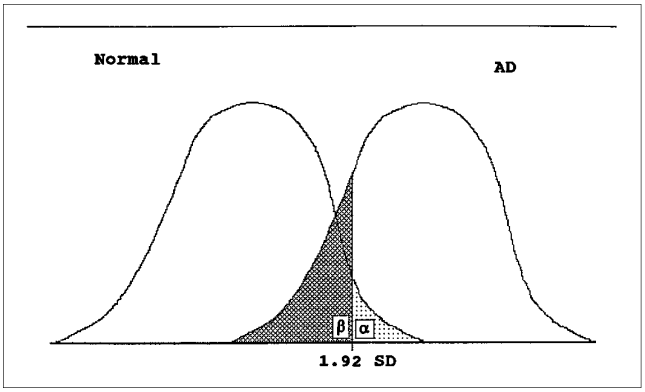

Fig 4. Calculation of sensitivity for a predetermined specificity of 95%. The probability of a false-positive value is α (specificity = $1 - \alpha$); β is the probability of a false-negative value (sensitivity = $1 - \beta$). α was chosen a priori 1.92 SD from the mean to equal a 95% probability of a true-positive value (95% specificity). The sensitivity of a measure is the area of the distribution of patients with Alzheimer disease outside of α as calculated by converting the z score to a probability estimate from a table of one-tailed standardized normal deviates (adapted from De Carli et al [5]).

TABLE 1: Clinical and demographic features of control subjects and patients with Alzheimer disease (AD)

	Control Subjects (n = 31)	Patients with Mild AD (n = 33)	P*	Patients with Moderate AD (n = 13)	P*
Women, n (proportion)	10 (0.32)	9 (0.27)	.87	2 (0.15)	.44
Age, y	69.1 ± 8.6	74.9 ± 8.0	.007	69.7 ± 9.8	.85
Education, y	8.2 ± 3.3	7.5 ± 4.4	.53	6.8 ± 3.5	.25
Apolipoprotein E ϵ4 allele, n (proportion)‡	9 (0.14)	25 (0.43)	.005	14 (0.54)	<.0005
MMSE score	29.0 ± 1.8	22.0 ± 2.1	<.0005	14.5 ± 1.7	<.0005
Basic ADL (functions lost)	0.00 ± 0.00	.53 ± 0.92	.002	.62 ± 1.12	.003
Instrumental ADL (functions lost)	0.0 ± 0.0	3.6 ± 2.1	<.0005	5.2 ± 2.8	<.0005
Disease duration, mo	· · ·	42.0 ± 27.8	· · ·	47.7 ± 29.3	· · ·
Clinical dementia rating	0.00 ± 0.00	0.95 ± 0.45	<.0005	1.96 ± 0.66	<.0005

Note.—Values are number (proportion) for sex and apolipoprotein E ϵ4 allele, and mean ± SD for all other variables. MMSE indicates Mini Mental State Examination; ADL, activities of daily living.

* Significance on χ^2 test (sex and apolipoprotein E) or *t* test (all other variables) versus control subjects.

‡ The total number of ϵ4 alleles in each group is twice the number of subjects in the group.

moval of variables from the model was based on F-to-remove values for λ lower than 2.71. The relevance of measures of hippocampal atrophy in the separation of patients from control subjects was tested in hypothesis-driven multivariable discriminant models in which measures of hippocampal atrophy were entered simultaneously.

The discriminant scores computed from each model were modeled into a gaussian curve as described for MoMs to compute sensitivity, with specificity set at 95%. Furthermore, sensitivity and specificity were addressed with Bayes's theorem (34), which minimizes overlapping between the two groups. The validity of each model was assessed by computing the sensitivity with which AD patients who were not used in the computation of the discriminant function (cross validation) were separated from control subjects.

Results

Clinical features of patients and control subjects are given in Table 1, which shows that both groups were similar in their sex distribution, age, and education. The prevalence of apolipoprotein E ϵ4 allele was greater in AD patients than in control subjects (0.42 versus 0.15; $\chi^2 = 12.1$; $df = 1$; $P = <.0005$), and was not different between mild and moderate AD (for prevalences, see Table 1; $\chi^2 = 1.3$; $df = 1$; $P = .25$). These figures are remarkably similar to those reported in recent studies of late-onset sporadic AD (35, 36), suggesting that our sample of AD patients and control subjects is representative of the general population.

Cognitive severity of the disease in AD patients was as expected on the basis of selection criteria: MMSE scores were much higher in patients with mild AD than in those with moderate AD; clinical dementia rating was around 1 (indicative of mild) and around 2 (indicative of moderate impairment) in the two groups. The pattern of preserved basic activities of daily living and mild to moderate loss of instrumental activities of daily living in the two groups was consistent with the cognitive severity.

Table 2 shows that, on average, all measures indicated greater atrophy in AD patients than in control subjects. Few measures (bifrontal index, interhemispheric fissure width, and minimum thickness of the right medial temporal lobe) were exceptions to this rule in patients with moderate AD in that statistical significance was not reached owing to a numerosity effect. Mean values in patients with mild and moderate AD were similar except for the width of the temporal horn, which showed a significant trend toward greater atrophy in the latter group (test for trend for right: $\beta = 1.48$; 95% CI, 0.91 to 2.05; $P < .00005$; test for trend for left: $\beta = 1.71$; 95% CI, 1.15 to 2.27; $P < .00005$; for smallest width of the temporal horn: $\beta = 1.95$; 95% CI, 1.39 to 2.51; $P < .00005$). The distribution of cranial width and measures of atrophy was normal (Komolgorov-Smirnov test, $P > .10$).

The relationship between atrophy and age was no different in men than in women for all measures except the width of the right temporal horn. The age times gender interaction ($\beta = -0.05$; 95% CI, −0.09 to −0.01; F(1,27) = 5.90; $P = .02$) and the age main effect ($\beta = 0.06$; 95% CI, 0.02 to 0.10; F(1,27) = 10.21; $P = .004$) for this measure in the general factorial ANOVA model indicated that the width of the right temporal horn increased with advancing

TABLE 2: Rough measures of cranial width and atrophy in control subjects and patients with Alzheimer disease (AD)*

	Control Subjects (n = 31)	Patients with Mild AD (n = 33)	P†	Patients with Moderate AD (n = 13)	P†
Cranial width	132.9 ± 6.9	129.5 ± 5.6	.04	134.4 ± 6.3	.53
Bifrontal index	29.0 ± 2.7	31.9 ± 2.6	<.0005	31.0 ± 5.4	.11
Interhemispheric fissure width	4.1 ± 1.5	5.4 ± 2.0	.007	5.0 ± 1.5	.09
Interuncal distance	26.8 ± 4.0	30.2 ± 3.8	.001	30.1 ± 4.9	.03
Minimum thickness of the medial temporal lobe					
R	14.5 ± 1.9	12.3 ± 3.1	.003	13.5 ± 1.7	.16
L	14.7 ± 1.6	12.4 ± 2.1	<.0005	10.6 ± 0.5	.02
Smallest‡	13.7 ± 1.5	11.4 ± 2.9	<.0005	10.6 ± 2.0	.002
Hippocampal height					
R	15.5 ± 1.6	13.8 ± 1.8	<.0005	13.8 ± 1.7	.003
L	14.7 ± 1.3	12.9 ± 1.9	<.0005	12.9 ± 3.2	.01
Smallest‡	14.3 ± 1.3	12.6 ± 1.7	<.0005	12.4 ± 2.6	.002
Width of the choroid fissure					
R	2.5 ± 1.1	4.5 ± 1.9	<.0005	3.9 ± 1.1	.001
L	2.8 ± 1.1	4.4 ± 1.5	<.0005	4.1 ± 1.7	.005
Largest‡	3.0 ± 1.2	4.8 ± 1.8	<.0005	4.5 ± 1.4	.001
Width of the temporal horn					
R	3.1 ± 1.3	5.0 ± 2.1	<.0005	6.0 ± 2.3	<.0005
L	3.3 ± 1.1	6.1 ± 2.3	<.0005	6.5 ± 1.9	<.0005
Largest‡	3.5 ± 1.2	6.5 ± 2.2	<.0005	7.2 ± 2.1	<.0005

* Measures are mean ± SD in millimeters.

† Significance on *t*-test versus control subjects.

‡ The right or left value indicating greater atrophy.

age in men and women, but more markedly in men.

Bifrontal index ($r = .47$; $P = .008$) and width of the right ($r = .46$; $P = .009$) and left ($r = .59$; $P = .001$) temporal horns were associated with age, whereas interuncal distance ($r = .59$; $P < .0001$) and width of the right ($r = .49$; $P = .005$) and left ($r = .47$; $P = .008$) temporal horns were associated with cranial width. Correlations of the other measures with age and cranial width were not significant and ranged from −0.02 to 0.31, and from −0.11 to 0.23, respectively. Since age is a known correlate of brain volume in healthy elderly subjects (37–39), MoMs were computed for all measures across values of age (ie, correcting the rough measure for the effect of greater atrophy with advancing age) (see "Subjects and Methods"). For interuncal distance and width of the temporal horn only, MoMs were computed also across values of cranial width. Furthermore, MoMs of the width of the right temporal horn were computed separately for men and women.

Table 3 shows the values of the measures of atrophy expressed in terms of MoMs and their expected sensitivity (with 95% specificity) for identifying patients with AD. When bilateral, the mean value of the measure indicating greater atrophy (left or right) is also reported. Overall, measures of global atrophy yielded lower sensitivity values than those of regional atrophy. The best sensitivity values were achieved by measurements of the width of the temporal horn for identifying patients with both mild and moderate AD, with the left measure proving more sensitive than the right. The distribution of rough values of the width of the largest temporal horn in patients and control subjects across age is shown in Figure 5.

Table 4 shows the results of the discriminant analyses. The discriminant function that was built on control subjects and patients with moderate AD with stepwise selection of variables selected two out of three measures of hippocampal atrophy and width of the interhemispheric fissure. Classification of cases on the basis of Bayes's theorem resulted in 96.8% sensitivity and 92.3% specificity (30 of 31 control subjects and 12 of 13 patients with moderate AD were classified correctly but are not reported in Table 4). The distribution of the discriminant scores of patients and control subjects was normal (Komolgorov-Smirnov test, $P > .10$), and gaussian modeling with specificity set at 95% resulted in a sensitivity value greater than 90% for moderate AD. Validation of

TABLE 3: Measures of atrophy expressed as multiples of the median in control subjects and patients with Alzheimer disease (AD)*

	Control Subjects (n = 31)	Cutoff for 95% Specificity	Patients with Mild AD (n = 33)	Sensitivity, %†	Patients with Moderate AD (n = 13)	Sensitivity, %†
Bifrontal index	1.00 ± 0.08	1.31	1.07 ± 0.11	29	1.06 ± 0.18	34
Interhemispheric fissure width	1.00 ± 0.38	1.63	1.28 ± 0.48	24	1.22 ± 0.39	15
Interuncal distance	1.00 ± 0.12	1.20	1.16 ± 0.15	40	1.10 ± 0.16	27
Minimum thickness of the medial temporal lobe						
R	1.00 ± 0.13	0.79	0.88 ± 0.21	33	0.93 ± 0.14	15
L	1.00 ± 0.10	0.84	0.86 ± 0.14	43	0.76 ± 0.23	63
Smallest‡	0.94 ± 0.09	0.79	0.80 ± 0.19	48	0.73 ± 0.15	66
Hippocampal height						
R	1.00 ± 0.10	0.84	0.88 ± 0.12	35	0.89 ± 0.11	31
L	1.00 ± 0.09	0.85	0.88 ± 0.13	41	0.88 ± 0.22	45
Smallest‡	0.96 ± 0.09	0.81	0.83 ± 0.11	43	0.82 ± 0.16	48
Width of the choroid fissure						
R	1.00 ± 0.44	1.73	1.69 ± 0.72	48	1.53 ± 0.37	30
L	1.00 ± 0.39	1.64	1.51 ± 0.50	39	1.46 ± 0.61	38
Largest‡	1.15 ± 0.45	1.89	1.77 ± 0.67	43	1.71 ± 0.50	36
Width of the temporal horn						
R	1.01 ± 0.37	1.62	1.62 ± 0.64	50	1.87 ± 0.74	63
L	1.00 ± 0.27	1.45	1.78 ± 0.67	69	1.96 ± 0.68	78
Largest‡	1.09 ± 0.29	1.57	2.00 ± 0.68	74	2.22 ± 0.72	82

* Measures are mean ± SD of multiples of the median (MoM), which are computed by regressing measures of atrophy on age and cranial width (interuncal distance and width of the temporal horn) or on age alone (all other measures).

† Sensitivity values are computed by gaussian modeling and specificity set at 95%.

‡ The right or left value of the MoM indicating greater atrophy.

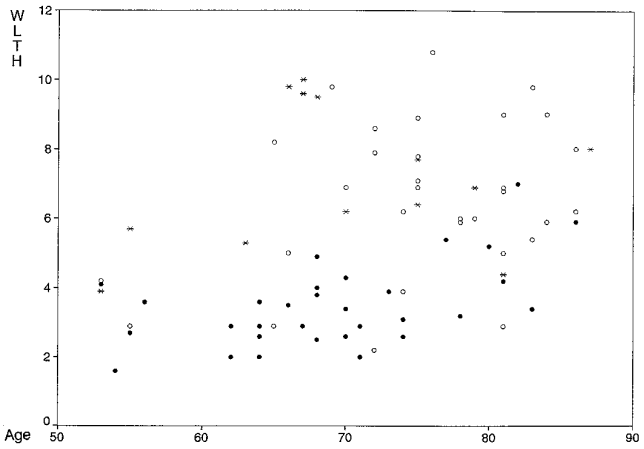

Fig 5. Measures of the width of the largest temporal horn in patients and control subjects in relation to age. *WLTH* indicates width of the largest temporal horn; *stars*, patients with moderate Alzheimer disease (AD) (n = 13); *open circles*, patients with mild AD (n = 33); and *solid circles*, control subjects (n = 31).

the discriminant model in patients with mild AD resulted in a sensitivity near 80%. The misclassified control subject was a 53-year-old woman with 12 years of education and a score of 30/30 on the MMSE. The *z* scores for age, education, and MMSE, computed on the basis of the value distribution of control subjects, were −1.88, 1.15, and 0.55, respectively. The misclassified patient with moderate AD was a 79-year-old woman with 5 years of education and a score of 13/30 on the MMSE. The *z* scores for age, education, and MMSE, computed on the basis of the value distribution of patients with moderate AD were −0.95, 0.53, and −0.85, respectively. These data indicate that neither misclassified subject was in the 5% tail distribution of their group for any of these variables and therefore could not be identified on clinical grounds.

The hypothesis-driven discriminant function that was built on control subjects and patients with moderate AD with simultaneous entering of measures of hippocampal atrophy showed sensitivity figures close to those of the previous model.

The discriminant function that was built on control subjects and patients with mild AD with stepwise selection of variables selected all three measures of hippocampal atrophy and interuncal distance. Classification of patients on the basis of Bayes's theorem resulted in 90.3% sensitivity and 81.8% specificity (28 of 31 control subjects and 27 of 33 patients with moderate AD were classified correctly but not reported in the Table). The distribution of the discriminant scores of patients and control subjects was nor-

TABLE 4: Discriminant functions and their cross validation in control subjects and patients with Alzheimer disease (AD)

	Discriminant Functions in Patients with Moderate AD*		Discriminant Functions in Patients with Mild AD*	
	Stepwise Selection	Hypothesis-driven	Stepwise Selection	Hypothesis-driven
Coefficients for				
Largest width of the temporal horn	2.14	1.96	1.03	1.33
Largest width of the choroid fissure	...	.29	0.99	0.90
Smallest hippocampal height	−4.76	−5.02	−4.17	−4.03
Interuncal distance	...	...	3.76	...
Width of the interhemispheric fissure	1.05	...	...	...
(Constant)	0.19	1.42	−3.24	0.20
Discriminant score				
Control subjects	−0.97 ± 0.86	−0.90 ± 0.74	−1.34 ± 0.84	−1.16 ± 0.65
Patients with mild AD	1.85 ± 1.75	1.67 ± 1.58	1.26 ± 1.13	1.09 ± 1.23
Patients with moderate AD	2.31 ± 1.29	2.15 ± 1.46	1.25 ± 1.35	1.38 ± 1.28
Sensitivity in the detection of				
Mild AD, %	79	80	86	83
Moderate AD, %	93	89	81	87

Note.—All variables are meant as multiples of the median; coefficients and the constant for computing discriminant scores are reported; discriminant scores are indicated as mean ± SD; and sensitivities are computed for a specificity figure of 95%.

* Discriminant functions in patients with moderate AD were computed for these patients and cross validated in patients with mild AD, and discriminant functions in patients with mild AD were computed for these patients and cross validated in patients with moderate AD. In the stepwise selection models, variables maximizing the distance between groups were selected, and in the hypothesis-driven models only measures of hippocampal atrophy were entered.

mal (Komolgorov-Smirnov test, $P > .10$), and gaussian modeling with specificity set at 95% resulted in a sensitivity of 86% for patients with mild AD. Validation of the discriminant model in patients with moderate AD resulted in a sensitivity around 80%. There was no difference for any of the clinical and demographic features between the 3 misclassified and the 28 correctly classified control subjects and between the 6 misclassified and the 27 correctly classifed patients with mild AD, again indicating that misclassified subjects could not be identified on clinical grounds. However, the discriminant score of the misclassified control subjects was closer to the control end of the discriminant score variable and significantly different from that of the correctly classified patients with mild AD (0.20 ± 0.23 versus 1.62 ± 0.90; t test, 2.70; $P = .01$). Similarly, the discriminant score of the misclassified patients with mild AD was closer to the AD end of the discriminant score variable and significantly different from that of the correctly classified control subjects (−0.39 ± 0.28 versus −1.50 ± 0.70; t test, 3.82; $P = .001$).

Once again, the hypothesis-driven discriminant function that was built on control subjects and patients with mild AD with simultaneous entering of measures of hippocampal atrophy showed sensitivity figures close to those of the previous model.

Discussion

We have shown that linear measures of hippocampal atrophy were able to differentiate patients with mild AD from healthy control subjects with a specificity of 95% and a sensitivity of around 85%. This held true after cross validation with a group of patients with moderate AD. Measures of atrophy outside the hippocampal region were not useful in differentiating patients with mild AD from control subjects.

Linear measures have been applied to improve the diagnostic reliability of AD (9, 10, 31, 32). It has been argued that linear measures, although potentially useful in the clinical routine, have poor diagnostic utility and yield an overall unsatisfactory sensitivity, and their sensitivity for discriminating patients with mild AD from healthy subjects has not been adequately addressed (40). We have shown that linear measurements can be taken with sufficient reliability to achieve sensitivity figures between 79% and 93%, and that these figures hold also in the mild stages of the disease.

Subjective ratings of atrophy have been reported to be reliable in differentiating patients with AD from control subjects (10, 11), to be useful in estimating severity of dementia (41), and to be superior to linear measures (40). However, it has recently been demonstrated that visual inspection of the medial temporal

lobe (hippocampus and hippocampal-amygdaloid complex) has poorer reliability for estimating temporal lobe atrophy than for assessing atrophy in other regions of the brain (42). This observation underlines the importance of easy-to-perform measurements of the medial temporal lobe in clinical practice.

Available evidence indicates that regional atrophy of the medial temporal lobe might be a sensitive indicator of the disease (7–18), even in its early stages (8). Therefore, it is not surprising that, also in the present study, variables independently discriminating patients with AD from healthy control subjects were those indicative of hippocampal and perihippocampal atrophy (width of the temporal horn, width of the choroid fissure, and hippocampal height). Only interuncal distance and width of the interhemispheric fissure significantly contributed to the discrimination of control subjects from patients with mild and moderate AD, respectively. This is in accordance with the meaning of these measures and the progression of the pathologic lesions over the course of the disease. In fact, interuncal distance has been proposed as a measure of atrophy of the medial temporal lobe (32, 43), although with conflicting results (44, 45), and the width of the interhemispheric fissure is an obvious marker of frontal lobe atrophy. Pathologic lesions in AD are known to appear early in the hippocampal region (46, 47), only later extending to the frontal lobe region (48). However, the exclusion of these measures from discriminant models, including indicators of hippocampal atrophy alone, only minimally affected sensitivity, suggesting that their practical contribution to the separation of AD patients from control subjects is trivial.

Previous studies have shown the size of the temporal horn to be the most sensitive indicator of the disease (12, 13, 15, 18, 19, 49). The present data in mildly demented patients confirm these observations in that the width of the temporal horn was the most sensitive single measure. The boundaries of the temporal horn at the level where we measured its width are represented by the hippocampus mesially and by the white matter containing axon fibers projecting to and from the hippocampal formation (hippocampus, entorhinal area, subiculum) laterally (46, 47, 50). Early lesions in AD heavily involve all these structures (46, 47, 51), giving rise to brain tissue atrophy (52). It can be hypothesized that loss of brain tissue around the temporal horns better reflects the combined involvement of hippocampal and perihippocampal structures occurring in patients with mild AD.

Hippocampal volume is presently the most sensitive measure for the anatomic diagnosis of AD in its mild stages (8, 53). In this study we have shown that a combination of linear measures of hippocampal atrophy separates patients with mild AD from control subjects with good sensitivity, and that adding other variables did not significantly increase sensitivity. We propose that the combination of linear measures of hippocampal atrophy that we identified with discriminant analysis can be considered a proxy of hippocampal volume.

Some potential drawbacks of the study should be considered. Patients with mild AD were significantly older than the control subjects. However, sensitivity and specificity of the measures of atrophy were addressed by transformation into MoMs. This implies taking the regression line of the atrophy measures on age in the control subjects as the reference for atrophy measures in the AD patients, thus removing the effect of normal aging on atrophy. The regression line could be an inappropriate reference for the patients if the age range of the control subjects is markedly narrower than that of the patients or if the age distribution of the control subjects is markedly skewed toward an extreme of the range. If these conditions are satisfied (which is true in our population of patients and control subjects, as shown in Fig 3), the age distribution of the patients cannot affect the precision of the MoM computation.

A few patients and control subjects were misclassified on the basis of measures of atrophy. On clinical grounds, patients with AD who were misclassified as control subjects were similar to correctly classified AD patients. The same held true for control subjects who were misclassified as AD patients, suggesting that clinical variables cannot enhance the likelihood of correct classification by means of measures of atrophy alone. On the other hand, misclassified control subjects yielded discriminant scores that could raise the possibility of misclassification. This observation constitutes evidence of validity of the discriminant score.

Pathologic confirmation was not available for our AD patients. However, clinical indications for the diagnosis of some degenerative non-AD dementias such as Pick disease (21), dementia

of the frontal type (22, 23), and dementia of the Lewy body type (24) have recently been provided. We believe that their exclusion from our AD sample increased the homogeneity of our AD group (54).

In conclusion, in the present work we have provided evidence of good sensitivity of linear measures of hippocampal atrophy for differentiating patients with mild AD from control subjects. We suggest that future work should focus on their specificity; that is, on the ability to discriminate AD from other degenerative and nondegenerative forms of dementia.

Acknowledgments

We thank Charles DeCarli and Ildebrando M. Appollonio for their worthy contribution to the discussion of the manuscript.

References

1. McKhann G, Drachman D, Folstein MF, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group. *Neurology* 1984;34:939–944
2. Kazee AM, Eskin TA, Lapham LW, Gabriel KR, McDaniel KD, Hamill RW. Clinicopathologic correlates in Alzheimer disease: assessment of clinical and pathologic diagnostic criteria. *Alz Dis Assoc Disord* 1993;3:152–164
3. Tierney MC, Fisher RH, Lewis AJ, et al. The NINCDS-ADRDA work group criteria for the clinical diagnosis of probable Alzheimer's disease: a clinicopathologic study of 57 cases. *Neurology* 1988;38:359–364
4. Winkler MA. Tacrine for Alzheimer's disease: which patient, what dose? *JAMA* 1994;271:1023–1024
5. DeCarli C, Kaye JA, Horowitz B, Rapoport S. Critical analysis of the use of computer-assisted transverse axial tomography to study human brain in aging and dementia of the Alzheimer type. *Neurology* 1990;40:872–883
6. Seab JP, Jagust WJ, Wong STS, Roos MS, Reed BR, Budinger TF. Quantitative NMR measurements of hippocampal atrophy in Alzheimer's disease. *Magn Reson Med* 1988;8:200–208
7. Kesslak JP, Nalcioglu O, Cotman CW. Quantification of magnetic resonance scans for hippocampal and parahippocampal atrophy in Alzheimer's disease. *Neurology* 1991;41:51–54
8. Jack CR, Petersen RC, O'Brien PC, Tangalos EG. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. *Neurology* 1992;42:183–188
9. Jobst KA, Smith AD, Szatmari M, et al. Detection in life of confirmed Alzheimer's disease using a simple measurement of medial temporal lobe atrophy by computed tomography. *Lancet* 1992; 340:1179–1183
10. Scheltens Ph, Leys D, Barkhof F, et al. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal aging: diagnostic value and neuropsychological correlates. *J Neurol Neurosurg Psychiatry* 1992;55:967–972
11. De Leon MJ, Golomb J, George AE, et al. The radiologic prediction of Alzheimer disease: the atrophic hippocampal formation. *AJNR Am J Neuroradiol* 1993;14:897–906
12. Killiany RJ, Moss MB, Albert MS, Sandor T, Tieman J, Jolesz F. Temporal lobe regions on magnetic resonance imaging identify patients with early Alzheimer's disease. *Arch Neurol* 1993;50: 949–954
13. Erkinjutti T, Lee DH, Gao F, et al. Temporal lobe atrophy on magnetic resonance imaging in the diagnosis of early Alzheimer's disease. *Arch Neurol* 1993;50:305–310
14. Cuénod C-A, Denys A, Michot J-L, et al. Amygdala atrophy in Alzheimer's disease: an in vivo magnetic resonance study. *Arch Neurol* 1993;50:941–945
15. Frisoni GB, Bianchetti A, Geroldi C, Trabucchi M, Beltramello A, Weiss C. Measures of medial temporal lobe atrophy in Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 1994;57:1438–1439
16. Ikeda M, Tanabe H, Nakagawa Y, et al. MRI-based quantitative assessment of the hippocampal region in very mild to moderate Alzheimer's disease. *Neuroradiology* 1994;36:7–10
17. Lehéricy S, Baulac M, Chiras J, et al. Amygdalohippocampal MR volume measurements in the early stages of Alzheimer disease. *AJNR Am J Neuroradiol* 1994;15:929–937
18. DeCarli C, Murphy DGM, McIntosh AR, Teichberg D, Schapiro MB, Horwitz B. Discriminant analysis of MRI measures as a method to determine the presence of dementia of the Alzheimer type. *Psychiat Res* 1995;57:119–130
19. Frisoni GB, Beltramello A, Weiss C, Geroldi C, Bianchetti A, Trabucchi M. Usefulness of simple measures of temporal lobe atrophy in probable Alzheimer's disease. *Dementia* 1996;7:15–22
20. DeCarli C, Maisog J, Murphy DGM, Teichberg D, Rapoport S, Horowitz B. Method for quantification of brain, ventricular, and subarachnoid CSF volumes from MR images. *J Comput Assist Tomogr* 1992;16:274–284
21. Knopman DS, Christensen KJ, Shut LJ, et al. The spectrum of imaging and neuropsychological findings in Pick's disease. *Neurology* 1989;39:362–368
22. Neary D, Snowden JS. Dementia of the frontal lobe type. In: Levin HS, Eisenberg HM, Benton AL, eds. *Frontal Lobe Function and Disfunction*. New York, NY: Oxford Press; 1991:304–317
23. Gustafson L. Clinical picture of frontal lobe degeneration of non-Alzheimer type. *Dementia* 1993;4:143–148
24. McKeith IG, Perry RH, Fairbairn AF, Jabeen S, Perry EK. Operational criteria for dementia of Lewy body type (SDLT). *Psychol Med* 1992;22:911–922
25. Folstein M, Folstein S, McHugh P. The Mini Mental State Examination. *J Psychiat Res* 1975;12:189–198
26. Hughes CP, Berg L, Danziger WL, Coben LA, Martin LA. A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982; 140:566–572
27. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. *Gerontologist* 1970;10:20–30
28. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9:179–186
29. Binetti G, Magni E, Padovani A, Cappa SF, Bianchetti A, Trabucchi M. Neuropsychological heterogeneity in mild Alzheimer's disease. *Dementia* 1993;4:321–326
30. Frisoni GB, Govoni S, Geroldi C, et al. Gene dose of the ϵ4 allele of apolipoprotein E and disease progression in sporadic late-onset Alzheimer's disease. *Ann Neurol* 1995;37:596–604
31. Barr AN, Heinze WJ, Dobben GD, Valvassori GE, Sugar O. Bicaudate index in computed tomography of Huntington's disease and cerebral atrophy. *Neurology* 1978;28:1196–1200
32. Dahlbeck SW, McCluney KW, Yeakley JW, Fenstermacher MJ, Bonmati C, Van Horn G. The interuncal distance: a new RM measurement for the hippocampal atrophy in Alzheimer's disease. *AJNR Am J Neuroradiol* 1991;12:931–932

33. Ebel RL. Estimation of the reliability of ratings. *Psychometrika* 1951;16:407–424
34. Norusis M. *Statistical Package for the Social Sciences,* rel. 5.0. Chicago, Ill: SPSS; 1992
35. Saunders AM, Strittmatter WJ, Schmechel D, et al. Association of apolipoprotein E allele ϵ4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 1993;43:1467–1472
36. Frisoni GB, Bianchetti A, Govoni S, et al. Association of apolipoprotein E E4 with vascular dementia. *JAMA* 1994;271:1317
37. Condon B, Grant R, Hadley D, Lawrence A. Brain and intracranial cavity volumes: in vivo determination by MRI. *Acta Neurol Scand* 1988;78:387–393
38. Jernigan TL, Press GA, Hesselink JR. Methods for measuring brain morphologic features on magnetic resonance images. *Arch Neurol* 1990;47:27–32
39. Coffey CE, Wilkinson WE, Parashos IA, et al. Quantitative cerebral anatomy of the aging human brain: a cross-sectional study using magnetic resonance imaging. *Neurology* 1992;42:527–536
40. Duara R. Neuroimaging with CT and MRI in Alzheimer disease. In: Terry RD, Katzman R, Bick KL, eds. *Alzheimer Disease.* New York: Raven Press, 1994:75–85
41. Launer LJ, Scheltens Ph, Lindeboom J, et al. Medial temporal lobe atrophy in an open population of very old persons: cognitive, brain atrophy, and socioeconomical correlates. *Neurology* 1995; 45:747–752
42. Victoroff J, Mack WJ, Grafton ST, Schreiber SS, Chui HC. A method to improve interrater reliability of visual inspection of brain MRI scans in dementia. *Neurology* 1994;44:2267–2276
43. Doraiswamy PM, McDonald WM, Patterson L, et al. Interuncal distance as a measure of hippocampal atrophy: normative data on axial MR. *AJNR Am J Neuroradiol* 1993;14:141–143
44. DeLeon MJ, Convit A, DeSanti S, et al. Value of interuncal distance measure in diagnosis of Alzheimer disease questioned. *AJNR Am J Neuroradiol* 1994;15:1286–1290
45. Early B, Escalona PR, Boyko OB, et al. Interuncal distance measurements in healthy volunteers and in patients with Alzheimer disease. *AJNR Am J Neuroradiol* 1993;14:907–910
46. Arnold SE, Hyman BT, Flory J, Damasio AR, Van Hoesen GW. The topographical and neuroanatomical distribution of neurofibrillary tangles and neuritic plaques in the cerebral cortex of patients with Alzheimer's disease. *Cereb Cortex* 1991;1:103–116
47. Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol* 1991;82:239–259
48. Hof PR, Morrison JH. The cellular basis of cortical disconnection in Alzheimer disease and related dementing conditions. In: Terry RD, Katzmann R, Bick KL, eds. *Alzheimer Disease.* New York: Raven Press, 1994:197–229
49. Wahlund LO, Andersson-Lundman G, Basun H, et al. Cognitive functions and brain structures: a quantitative study of CSF volumes on Alzheimer patients and healthy control subjects. *Magn Reson Imaging* 1993;11:169–174
50. De Armond SJ, Fusco MM, Dewey MM. *Structure of the Human Brain: A Photographical Atlas.* New York, NY: Oxford University Press; 1989
51. Hof PR, Bierer LM, Perl DP, et al. Evidence for early vulnerability of the medial and inferior aspects of the temporal lobe in an 82-year-old patient with preclinical signs of dementia: regional and laminar distribution of neurofibrillary tangles and senile plaques. *Arch Neurol* 1992;49:946–953
52. Huesgen CT, Burger PC, Crain BJ, Johnson GA. In vitro MR microscopy of the hippocampus in Alzheimer disease. *Neurology* 1993;43:145–152
53. Bhatia S, Bookheimer SY, Gaillard WD, Theodore WH. Measurement of whole temporal lobe and hippocampus for MR volumetry: normative data. *Neurology* 1993;43:2006–2010
54. Frisoni GB, Trabucchi M, Pizzolato G. Frontal lobe dementia. *Int J Geriatr Psychiatry* 1993;8:357