Atrophy of medial temporal lobes on MRI in “probable” Alzheimer’s disease and normal ageing: diagnostic value and neuropsychological correlates

Ph Scheltens, D Leys, F Barkhof, D Huglo, H C Weinstein, P Vermersch, M Kuiper, M Steinling, E Ch Wolters, J Valk

Abstract

Magnetic resonance imaging (MRI) has shown a great reduction in medial temporal lobe and hippocampal volume of patients with Alzheimer’s disease as compared to controls. Quantitative volumetric measurements are not yet available for routine clinical use. We investigated whether visual assessment of medial temporal lobe atrophy (MTA) on plain MRI films could distinguish patients with Alzheimer’s disease (n = 21) from age matched controls (n = 21). The degree of MTA was ascertained with a rating procedure and validated by linear measurements of the medial temporal lobe including the hippocampal formation and surrounding spaces occupied by cerebrospinal fluid.

Patients with Alzheimer’s disease showed a significantly higher degree of subjectively assessed MTA than controls (p = 0.0005). Linear measurements correlated highly with subjective assessment of MTA and also showed significant differences between groups. Ventricular indices did not differ significantly between groups. In Alzheimer’s disease patients the degree of MTA correlated significantly with scores on the mini-mental state examination and memory tests, but poorly with mental speed tests. This study shows that MTA may be assessed quickly and easily with plain MRI films. MTA shown on MRI strongly supports the clinical diagnosis of Alzheimer’s disease, is related to memory function, and seems to occur earlier in the disease process than does generalised brain atrophy.

(J Neurol Neurosurg Psychiatry 1992;55:967–972)

One of the first and most striking manifestations of Alzheimer’s disease is memory impairment. The hippocampus and adjacent structures are crucial to memory; their bilateral destruction leads invariably to the loss of this essential intellectual faculty. Ball et al have stated that Alzheimer’s disease is mainly a “hippocampal dementia” and have shown nerve cell loss and gliosis in both hippocampi of diseased patients while other brain regions were unaffected. Other pathological studies have also incriminated the temporal basal gray matter structures, including the hippocampus, as the brain areas affected earliest and most severely by Alzheimer’s disease.⁵

Recent neuroimaging studies have also focused attention on these structures. A computed tomography (CT) study by de Leon and coworkers reported a significant widening of the hippocampal fissure in Alzheimer’s disease patients.⁶ CT imaging of the temporal lobes, however, is hampered by bone hardening artifact and a limited viewing angle; it shows enlargement of the temporal horns and widening of the hippocampal fissures as indirect changes due to atrophy of the hippocampal complex. Magnetic resonance imaging (MRI) yields superior visualisation of the medial temporal lobe, including the hippocampus.⁷ It also allows an imaging plane perpendicular to the long axis of the hippocampus, thus reducing volume averaging. MRI technique permits direct visualisation of the hippocampal formation in substantial cytoarchitectonic detail.⁸

Two studies have shown over 40% reduction in hippocampal volume in Alzheimer’s disease patients as compared to controls.⁹¹⁰ Quantitative volumetric measurements, however, are not yet available for routine clinical use and therefore we investigated whether a visual assessment of medial temporal lobe atrophy (MTA) on plain MRI films could distinguish Alzheimer’s disease patients from controls. Validation of visually assessed atrophy was attempted by linear measurements of the medial temporal lobe. For comparison we obtained two indices of general ventricular enlargement to study their ability to separate groups. We also studied whether medial temporal lobe atrophy correlated with neuropsychological impairment in patients with “probable” Alzheimer’s disease.

Subjects and methods

Patients

This study was carried out in Amsterdam, Netherlands, and in Lille, France. Forty two right handed subjects (21 Alzheimer’s disease, 21 controls) were included: 21 were French speaking (11 Alzheimer’s disease, 10 controls) and 21 were Dutch speaking (10 Alzheimer’s disease, and 11 controls). The group with Alzheimer’s disease included 19 women and two men, mean (SD) age 72.8 (11.5) years (range 52–90 years), with a mean mini-mental...
state examination score of 14-9 (6-17) (range 1-24). The diagnosis of "probable" Alzheimer's disease was made according to current research criteria. We excluded patients with evidence of neurological disease other than Alzheimer's disease, incurring toxic, metabolic, traumatic, demyelinating, neoplastic, or focal cerebrovascular disease. Each subject had a Hachinski score of 3 or less. None of the patients suffered from arterial hypertension, diabetes mellitus, or cardiac disease, according to our previously published criteria. Nine patients had presenile Alzheimer's disease (onset before 65th year) and 12 had senile Alzheimer's disease. None of them had a definite family history of the disease but two had a first degree relative suffering from dementia, without sufficient data available to make a diagnosis of Alzheimer's disease. None of the patients displayed myoclonus or extrapyramidal signs.

Controls
Control subjects were healthy volunteers. They were spouses of patients or volunteers from a general practice. These subjects (8 women, 13 men; mean age 70-9 (10-6; 54-94) years) had no history of mental decline or strokes and had a mini-mental state examination score higher than 27 (mean 29-7 (0-56)). Each of them had given informed consent after the nature of the procedure had been fully explained.

### Table 1

<table>
<thead>
<tr>
<th>Score</th>
<th>Width of choroid fissure</th>
<th>Width of temporal horn</th>
<th>Height of hippocampal formation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>1</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>2</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>3</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>4</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
</tbody>
</table>

† = increase, ‡ = decrease. N = normal.

### MRI technique

MRI was performed on an MR-max (General Electric) with a superconducting magnet operating at a field strength of 0-5 tesla, or on a Teslacore II (Technicare) operating at 0-6 tesla. T1 weighted spin echo sequences were used (TR 300 ms, TE 22 ms, six sequences, field of view 20). Six oblique slices (slice thickness 5 mm, interslice gap 1 mm, in plane resolution 0-8 x 1-0 mm) parallel to the brainstem axis were planned from a midsagittal scout image (fig 1). In both centres the first image was planned directly adjacent to the brainstem, resulting in identical images for all subjects. Total acquisition time was six minutes.

### MRI analysis

MTA was assessed visually by five of us who were blinded to diagnosis and age of the subjects. A definite score was assigned when consensus was reached. The procedure was performed as follows. All 42 scans were ranked from absent to severe atrophy, judged on all six slices. The scans were then pooled in groups with various degrees of atrophy, resulting in five different scores (0-4), subsequently referred to as subjective MTA score. The characteristics of each group were later defined in terms of height of the hippocampal formation (hippocampus and parahippocampal gyrus) and enlargement of the surrounding cerebrospinal fluid spaces, for application by others (table 1). Example are shown in figure 2. In addition, four linear measures of the medial temporal lobe (A, B, C, D) were calculated on both sides (left and right). A was the greatest height of the hippocampal formation, defined as the dentate gyrus, hippocampus proper, and subiculum together with the parahippocampal gyrus; B was the greatest...
width between the hippocampal formation and
the brainstem; C was the vertical width of the
choroid fissure centred on the midpoint of the
hippocampus; and D was the width of the
temporal horn (fig 3). These measures were
divided by the width between the inner tables
of the calvarium on the same level, designated
as the brain width, resulting in four linear
indices (AI, BI, CI, and DI) for each side.

For comparison we measured the lateral
ventricular index and the third ventricle index
on the same slice used for the temporal indices.
The ventricular index (VI) was defined as the
width between the lateral ventricles divided by
the width between the inner tables of the
calvarium on the same level and the third
ventricle index (V3I) was defined as the width
of the third ventricle divided by the width
between the inner tables of the calvarium on
the same level. This method is essentially
the same as described for CT.15

All linear measurements were taken on the
slice that best depicted both hippocampal
formations, usually the fourth slice, after mag-
nification of the hard copy of the scan by
projecting it on a screen with an overhead
projector to five times real size. The measure-
ments were later divided by five, providing
actual sizes. Linear measurements were per-
formed by two neuroradiologists together in
one session.

Neuropsychological assessment
Neuropsychological tests were performed in
Alzheimer's disease patients by an investigator
whose native language was the same as that of
the patient. Tests were standardised in both
languages. They consisted of tests for memory
and mental speed.

Memory was evaluated by the Fuld object
memory evaluation,16 Wechsler logical mem-
ory test17 for free recall, the digit span18 for
immediate verbal memory, and word fluency18
for semantic memory (categorical names of
foods). The object memory evaluation con-
isted of a "tactile" part in which 10 common
objects in a bag were presented to patients to
determine whether they could identify objects
by touch (stereognosis), providing a "tactile"
score. After naming or describing the object
the patient took the object out to "see if (s)he
was right" and to name it if it had not
been recognised by touch, providing a "visual"
score. The scores for tactile and visual recogni-
tion were summed. Next the patient was asked
to recall the objects from the bag, and was
given four more chances to learn and recall
them. Recall was tested 15 minutes later and a
recall score obtained.

Mental speed was assessed by the trail making
test, part A,18 and the digit symbol substitution
test.19 The trail making and word fluency tests
were performed during the 15 minute period
between the fifth trial and the delayed recall of
the object memory evaluation. The mini-
mental state examination was used as a meas-
ure of the severity of dementia in Alzheimer's
disease patients.

Statistical analysis
Groups were compared using means and 95%
confidence intervals obtained by a t test.
Ordinal variables were compared with the
Mann Whitney U test. Correlations were cal-
culated with Spearman's non-parametric rank
correlation test (r'), with Bonferroni's correc-
tion for multiple comparisons when appro-
priate. p Values < 0.05 were regarded as
significant. Null hypotheses were tested two
sided. Data were analysed with the SPSS/PC
statistical package.

Results
The groups did not differ significantly with
respect to mean age and mean brain width. In
the control group the brain width and the
subjective score were not significantly different
between men and women.

The breakdown of the subjective MTA
scores is given in figure 4. Seventeen out of 21
(81%) patients had scores ranging from 2 to 4,
while 14/21 (67%) controls had scores of 0 or
1. None of the controls had a score of 4, and
one patient had a score of 0 (further details of
this patient are given in the discussion). The
difference between groups was significant
(p = 0.0005). The subjective score did not

---

Figure 3 Schematic drawing showing linear measures of medial temporal lobe. A = largest vertical height of hippocampal formation, defined as dentate gyrus, hippocampus proper, and subiculum together with parahippocampal gyrus; B = largest horizontal width between hippocampal formation and brainstem; C = largest vertical width of choroid fissure; D = width of temporal horn.

---

Figure 4 Breakdown of subjective scores of medial temporal lobe atrophy in Alzheimer's disease patients (solid bars) and controls (hatched bars). Difference is significant (p = 0.0005, Mann Whitney U test).
differ between patients with presenile and senile Alzheimer's disease.

The linear indices are expressed as means with 95% confidence intervals in Table 2. They differed significantly between groups, except for left and right BI and for both indices of general ventricular enlargement. The linear indices of the medial temporal lobe, except left and right BI, correlated significantly with the subjective MTA score \( r > 0.60; p < 0.05 \).

The mini-mental state examination score correlated significantly with the subjective MTA score, left and right AI, and left DI in Alzheimer's disease patients \( (n = 21) \). Further neuropsychological testing could be performed in only 19 patients because of severe aphasia in two patients. Correlations of neuropsychological parameters in patients and the subjective MTA score and the linear indices, without BI, are shown in Table 3. Because of multiple comparisons (56 correlations) we applied a Bonferroni correction yielding a \( p \) value of 0.05/56 = 0.0009—that is, \( r > 0.70 \) (Table 3).

The left CI, right CI and right DI did not correlate significantly with any of the neuropsychological parameters. Of the memory tests, object memory after 15 minutes correlated significantly with the subjective score as well as with left AI and left DI.

**Discussion**

Using subjective visual assessment we found significant differences in medial temporal lobe atrophy between patients with Alzheimer's disease and age matched controls. This method correctly identified 81% of Alzheimer's disease patients (sensitivity) and 67% of controls (specificity). The subjective MTA score correlated well with linear measurements of the medial temporal lobe, which also showed significant differences between groups. Good correlations between visual assessment and linear measurements have been shown earlier in studies of ventricular and sulcal enlargement on CT\(^1\) and MRI.\(^2\)

Both types of assessment of MTA did not completely separate Alzheimer's disease patients from controls. The overlap may be explained either by the presence of MTA in a small portion of the controls as an indicator of preclinical Alzheimer's disease or, more likely, by an age related hippocampal atrophy. To elaborate on this further we performed a linear regression analysis with the subjective MTA score as dependent variable. This analysis showed that the subjective score depended significantly on diagnosis (Alzheimer's disease or control) and also on age, and not on sex (see Table 4). Post hoc analysis for the two groups separately (Fig 5) found a correlation with age in the control group \( (r = 0.59, p = 0.066) \), but not in the Alzheimer's disease group \( (r = 0.37, p = 0.10) \), suggesting a poor sensitivity for the diagnosis of Alzheimer's disease in older patients. From Figure 5 it may be inferred that of the seven controls with scores 2 or 3, six

---

**Table 2** Mean (95% CI) of linear indices for patients with Alzheimer's disease and controls

<table>
<thead>
<tr>
<th></th>
<th>Alzheimer's disease patients</th>
<th>Controls</th>
<th>( p ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: left</td>
<td>0.155 (0.138 to 0.172)</td>
<td>0.182 (0.169 to 0.194)</td>
<td>0.02</td>
</tr>
<tr>
<td>left</td>
<td>0.166 (0.153 to 0.179)</td>
<td>0.192 (0.181 to 0.203)</td>
<td>0.007</td>
</tr>
<tr>
<td>B: left</td>
<td>0.099 (0.046 to 0.072)</td>
<td>0.056 (0.046 to 0.066)</td>
<td>NS</td>
</tr>
<tr>
<td>right</td>
<td>0.062 (0.051 to 0.073)</td>
<td>0.099 (0.050 to 0.068)</td>
<td>NS</td>
</tr>
<tr>
<td>C: left</td>
<td>0.035 (0.025 to 0.045)</td>
<td>0.019 (0.014 to 0.024)</td>
<td>0.007</td>
</tr>
<tr>
<td>right</td>
<td>0.034 (0.022 to 0.046)</td>
<td>0.014 (0.010 to 0.017)</td>
<td>0.004</td>
</tr>
<tr>
<td>D: left</td>
<td>0.070 (0.054 to 0.087)</td>
<td>0.035 (0.025 to 0.045)</td>
<td>0.001</td>
</tr>
<tr>
<td>right</td>
<td>0.056 (0.042 to 0.069)</td>
<td>0.027 (0.018 to 0.037)</td>
<td>0.001</td>
</tr>
<tr>
<td>V3</td>
<td>0.061 (0.051 to 0.069)</td>
<td>0.061 (0.051 to 0.069)</td>
<td>NS</td>
</tr>
<tr>
<td>VI</td>
<td>0.310 (0.297 to 0.323)</td>
<td>0.300 (0.287 to 0.313)</td>
<td>NS</td>
</tr>
</tbody>
</table>

Indices obtained by dividing the following variables by brain width: A, greatest height of hippocampal formation; B, greatest width between hippocampal formation and brainstem; C, width of choroid fissure; D, width of temporal horn; V3, width of third ventricle; VI, width between lateral ventricles.

---

**Table 3** Spearman rank correlations between neuropsychological data and medial temporal lobe atrophy in “probable” Alzheimer's disease patients \( (n = 19) \)

<table>
<thead>
<tr>
<th></th>
<th>Subjective score</th>
<th>AI Left</th>
<th>Right</th>
<th>CI Left</th>
<th>Right</th>
<th>DI Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object memory evaluation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tactile</td>
<td>-0.07</td>
<td>0.24</td>
<td>0.16</td>
<td>-0.02</td>
<td>-0.23</td>
<td>-0.32</td>
<td>0.11</td>
</tr>
<tr>
<td>Sum</td>
<td>-0.29</td>
<td>0.35</td>
<td>0.44</td>
<td>-0.17</td>
<td>-0.35</td>
<td>-0.32</td>
<td>0.21</td>
</tr>
<tr>
<td>Recall after 15 minutes</td>
<td>-0.79*</td>
<td>0.71*</td>
<td>0.49</td>
<td>-0.30</td>
<td>-0.15</td>
<td>-0.04</td>
<td>-0.36</td>
</tr>
<tr>
<td>Wechsler logical memory</td>
<td>-0.52</td>
<td>0.49</td>
<td>0.34</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.48</td>
<td>-0.13</td>
</tr>
<tr>
<td>Digit span(^1)</td>
<td>-0.25</td>
<td>0.27</td>
<td>0.28</td>
<td>-0.17</td>
<td>-0.23</td>
<td>-0.41</td>
<td>0.17</td>
</tr>
<tr>
<td>Word fluency</td>
<td>-0.17</td>
<td>0.42</td>
<td>0.14</td>
<td>-0.17</td>
<td>-0.14</td>
<td>-0.50</td>
<td>0.14</td>
</tr>
<tr>
<td>Mental speed:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trail making test</td>
<td>-0.32</td>
<td>-0.39</td>
<td>0.33</td>
<td>-0.11</td>
<td>-0.28</td>
<td>-0.40</td>
<td>0.21</td>
</tr>
<tr>
<td>Digit symbol substitution test</td>
<td>-0.43</td>
<td>-0.48</td>
<td>0.21</td>
<td>-0.27</td>
<td>-0.24</td>
<td>-0.47</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

\(^*\)Significant correlations after Bonferroni correction \( r > 0.70 \).

\(^1\)Total of digits forward and backward.
were aged over 77. In the CT study of de Leon et al sensitivity dropped greatly after the age of 60. The MRI studies of Seab et al and Kesslak et al found no overlap. This may be explained by several factors that differed from our study. Both studied small groups, consisting of 10 patients and seven controls in Seab et al’s study, and eight patients and seven controls in that of Kesslak et al. In such small groups an overlap is less likely to occur. In addition, the mean age of the controls was lower than that of the Alzheimer’s disease patients and the oldest patient was 78 years in the former study and 80 years in the latter. Our study not only included more patients but the age range was also wider. The two previous studies used volumetric measures involving only the hippocampus or both the hippocampus and parahippocampal gyrus. Although this may be a better way to characterise the medial temporal lobe, it is unlikely that it completely separates age matched controls from Alzheimer’s disease patients in larger groups.

Only one patient exhibited a subjective score of 0. This was a left handed woman of 61 years with a mini-mental state examination score of 16. She suffered from mild aphasia together with asterognosis of the left hand and visuospatial disturbances. She performed reasonably well on the Wechsler logical memory and object memory evaluation tests, illustrating the relatively preserved memory function. One year later her memory functions had worsened. This patient resembles an earlier reported (present) 2 Alzheimer’s disease patient who presented with focal symptoms instead of definite memory disturbances.

Differences in subjective scores might be due to the preponderance of women in the Alzheimer’s disease group, since brain size may be smaller in women. Although this could have influenced the ranking procedure we think that it is not of major importance since brain width did not differ significantly between groups. Furthermore, within the controls no differences in brain width or subjective score were found between men and women, as in previous reports, and multiple linear regression analysis did not identify sex as an explanatory variable of the subjective score.

In contrast to earlier reports, linear indices of the lateral ventricles and third ventricle as indicators of generalised brain atrophy showed no significant differences between the two groups in this study. These indices are well known to correlate with severity of dementia and are of more value in longitudinal studies than in cross sectional studies. The failure to find significant differences between groups in our study is probably due to small sample size. It does illustrate, however, that in spite of the few subjects MTA differed significantly between groups, and does not (only) reflect overall brain atrophy in Alzheimer’s disease. MTA seems to occur early in the disease process and might be a selective marker for Alzheimer’s disease. The strong correlations of the delayed recall in the object memory evaluation test with some of the MTA variables is striking. Some of the MTA variables also correlated significantly with the mini-mental state examination score. This does not imply, in our opinion, that the correlations with this memory test are just the result of global dementia, since significant correlations with mental speed tests would then also be present. Our findings strongly support the hippocampal model of memory and emphasise the potency of MTA for early diagnosis.

From these data we cannot decide which type of assessment of MTA is to be preferred. Linear measurements in general are subject to greater interobserver variability, may vary with the MRI technique, and rely highly on good scan quality (without motion artifacts). Subjective visual assessment seems an easy and quick method. It has been shown to correlate well with computer assisted planimetric and volumetric measurements.

We think that subjective assessment of medial temporal lobe atrophy may be useful to support a clinical diagnosis of “probable” Alzheimer’s disease, enabling the clinician to differentiate this diagnosis from normal aging and “pseudo” dementia—that is, major depression—in subjects younger than 75 years. To date no other types of dementia have been studied with respect to MTA on MRI; therefore the value of assessment of MTA in the differential diagnosis of dementia still has to be established.

9 Jack CR Jr, Wouncy CK, Zinemeister AR, Sharbrough FW, Petersen RC, Cascino GD. Anterior temporal lobes and


