Visual Assessment of Atrophy on Magnetic Resonance Imaging in the Diagnosis of Pathologically Confirmed Young-Onset Dementias

Marcus Likeman, FRCP; Valerie M. Anderson, BSc; John M. Stevens, DRACR, FRCP; Adam D. Waldman, PhD, MRCP, FRCP; Alison K. Godbolt, MA, MRCP; Chris Frost, MA, DipStat; Martin N. Rossor, MD, FRCP; Nick C. Fox, MD, FRCP

**Objectives:** To investigate the diagnostic accuracy of visual inspection of magnetic resonance imaging (MRI) in a range of pathologically confirmed diseases causing young-onset dementia and to assess the sensitivity and specificity of atrophy patterns for Alzheimer disease (AD) and frontotemporal lobar degeneration (FTLD).

**Design:** Sixty-two patients with pathologically confirmed diseases that may present as young-onset dementia were selected from a biopsy and postmortem series. The first diagnostic T1-weighted volumetric MRI was obtained for each patient, together with images from 22 healthy control subjects. All MRIs were assessed for regional atrophy independently by 3 neuroradiologists, blinded to all clinical details except age. Observers were also asked to use their clinical judgment to form a diagnosis.

**Results:** Eighty-seven percent of dementia cases were distinguished from controls after visual inspection of MRI, and a correct pathologically confirmed diagnosis was given in 58% of cases. Hippocampal atrophy was noted in 92% of AD cases but was commonly seen in other dementias and controls. A bilateral symmetrical pattern of hippocampal atrophy discriminated AD from FTLD with 47% specificity, while posterior greater than anterior gradient of atrophy was 92% specific for AD. Atrophy of the anterior, inferior, and lateral temporal lobes was suggestive of FTLD pathology (≥90% sensitivity), while anterior greater than posterior gradient of atrophy and hemispheric asymmetry of atrophy were each at least 85% specific for FTLD.

**Conclusion:** Despite variation and overlap of atrophy patterns, visual inspection of regional atrophy on MRI may aid in discriminating AD and FTLD.

*Arch Neurol. 2005;62:1410-1415*
Table 1. Demographic and Clinical Characteristics of Subjects at the Time of Magnetic Resonance Imaging

<table>
<thead>
<tr>
<th>Group</th>
<th>Total No. (No. of Men)</th>
<th>Age, y</th>
<th>MMSE Score</th>
<th>Disease Duration, y</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>25 (13)</td>
<td>59 (10.4)</td>
<td>15 (5.0)†‡</td>
<td>4.7 (3.5)</td>
</tr>
<tr>
<td>FTLD</td>
<td>17 (11)</td>
<td>56 (10.1)</td>
<td>22 (7.1)†</td>
<td>3.5 (2.0)</td>
</tr>
<tr>
<td>Other pathologies</td>
<td>20 (15)</td>
<td>58 (14.0)</td>
<td>20 (6.0)†</td>
<td>3.8 (4.1)</td>
</tr>
<tr>
<td>Control</td>
<td>22 (11)</td>
<td>53 (10.8)</td>
<td>30 (0.8)</td>
<td>...</td>
</tr>
</tbody>
</table>

Abbreviations: AD, Alzheimer disease; FTLD, frontotemporal lobar degeneration; MMSE, Mini-Mental State Examination; ellipsis, not applicable.

*Data are given as mean (SD) unless otherwise indicated.
†Significantly worse score on MMSE than control group (P<.001).
‡Significantly worse score on MMSE than FTLD and other pathologies groups (P<.05).

MAGNETIC RESONANCE IMAGING

Patients with dementia had T1-weighted volumetric MRI performed as part of a diagnostic workup when they presented with cognitive complaints to the National Hospital for Neurology and Neurosurgery. Control subjects had T1-weighted volumetric imaging performed as part of longitudinal research studies. Images were acquired on 1.5-T MRI scanners (Horizon Echospeed, version 3.7 or LX; GE Medical Systems, Milwaukee, Wis) as part of clinical imaging protocols. All were acquired in the coronal plane using spoiled gradient echo techniques (spoiled gradient recalled echo and magnetization-prepared rapid gradient echo) and single excitation using a 24-cm field of view to give 124 contiguous 1.5-mm sections. Other acquisition parameters varied depending on the scanner and clinical protocol used, with a repetition time between 13 and 35 milliseconds, an echo time between 4 and 9 milliseconds, and a matrix size of 256 × 256, 200 × 200, or 180 × 180 pixels.

MRI ANALYSIS

Magnetic resonance images were viewed on a Sun workstation (Sun Microsystems, Inc, Santa Clara, Calif) using software that allows images to be viewed in 3 orthogonal planes.

Images were assessed for atrophy in random order and independently by 3 experienced neuroradiologists (M.L., J.M.S., and A.D.W.) who were blinded to all clinical details except age at the time of MRI. To assess intraobserver reliability, 15 cases (chosen arbitrarily by a fourth observer [V.M.A.]) appeared twice in the series, making a total of 99 scan assessments for each observer. Only the first assessment of these 15 cases was included in the pathological-radiological analysis. Assessments of whole-brain gradient (anterior-posterior) and asymmetry (left-right hemispheric) of atrophy were made. The frontal lobes were assessed for the presence or absence of atrophy, while 4 regions of the left and right temporal lobes were rated using a modification of the rating scales by Scheltens and Galton and their colleagues. Hippocampal atrophy was rated when visualized in the coronal plane at the level of the hippocampal head on a scale of 0 to 4 (0, none; 1, minimal; 2, mild; 3, moderate; and 4, severe). Anterior temporal lobe (ATL) or amygdala, lateral temporal gyri (LTG) or fusiform gyrus, and parahippocampal gyrus atrophy were assessed on a scale of 0 to 3 (0, none; 1, mild; 2, moderate; and 3, severe). Each observer formed a diagnosis for each patient, based on regional atrophy patterns. Observers knew the possible diagnoses but did not know the distribution of cases among them. A category of “non-specific changes” was included for imaging results that were abnormal but that could not be placed in a specific diagnostic category. The FTLD category consisted of several associated pathologic substrates, including Pick disease, dementia lacking distinctive histology, and dementia with ubiquitin-
RESULTS

Groups were matched for age and disease duration. Mini-Mental State Examination scores were significantly reduced in dementia groups compared with controls, and in AD compared with FTLD and other dementias (Table 1). Although image quality varied, all images allowed adequate visual assessment of atrophy. The mean (SD) intraobserver agreement of radiological diagnoses was 84% (7.7%), and interobserver agreement (between pairs of observers) based on all 84 subjects in the study was 63% (6.6%). Correcting for chance agreement between observers, $\kappa$ was 50% ($P<.001$). Among all subjects, a mean (SD) of 87% (7.3%) of dementia cases were distinguished from controls on visual assessment of MRIs; pathologically confirmed diagnoses of 58% (3.6%) were correctly identified from imaging.

Of 25 cases of primary AD, a mean (SD) of 84% (10.6%) showed abnormal findings, and 67% (8.3%) were correctly diagnosed as AD, distinct from those with other dementias and controls. Of misdiagnosed AD cases, 17% were thought to be normal, 9% FTLD, and 7% other pathologies. Table 2 lists the sensitivity and specificity for visual assessment of regional atrophy in pathologically confirmed AD and FTLD. Hippocampal atrophy (left or right, any severity) was 92% sensitive for AD. However, specificity was 62% when AD was compared with controls and decreased to 6% when compared with FTLD. Specificity improved when moderate to severe (left or right) hippocampal atrophy (atrophy graded as $\geq$3) was considered, but sensitivity decreased to 41%. Bilateral sym-

### Table 2. Mean Sensitivity and Specificity of Regional Atrophy, Assessed on Magnetic Resonance Imaging by 3 Neuroradiologists, for AD and FTLD Pathologies

<table>
<thead>
<tr>
<th>Pathology</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>vs Control</td>
<td>vs All Other Cases</td>
</tr>
<tr>
<td><strong>AD</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hippocampal atrophy</td>
<td>92</td>
<td>62</td>
</tr>
<tr>
<td>Any severity</td>
<td>92</td>
<td>61</td>
</tr>
<tr>
<td>Moderate to severe</td>
<td>92</td>
<td>57</td>
</tr>
<tr>
<td>Bilateral and symmetrical</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td>Left to right hemispheric asymmetry</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td>Posterior greater than anterior gradient of atrophy</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td>Bilateral and symmetrical hippocampal atrophy</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td>or posterior greater than anterior gradient of atrophy</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td><strong>FTLD</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hippocampal atrophy (any severity)</td>
<td>94</td>
<td>...</td>
</tr>
<tr>
<td>Parahippocampal gyrus atrophy (any severity)</td>
<td>92</td>
<td>...</td>
</tr>
<tr>
<td>ATL or amygdala atrophy (any severity)</td>
<td>92</td>
<td>...</td>
</tr>
<tr>
<td>LGT or fusiform gyrus atrophy (any severity)</td>
<td>90</td>
<td>...</td>
</tr>
<tr>
<td>Parahippocampal gyrus atrophy (moderate-severe)</td>
<td>65</td>
<td>...</td>
</tr>
<tr>
<td>ATL or amygdala atrophy (moderate-severe)</td>
<td>69</td>
<td>...</td>
</tr>
<tr>
<td>LGT or fusiform gyrus atrophy (moderate-severe)</td>
<td>55</td>
<td>...</td>
</tr>
<tr>
<td>Frontal lobe atrophy</td>
<td>76</td>
<td>...</td>
</tr>
<tr>
<td>Anterior greater than posterior gradient of atrophy</td>
<td>39</td>
<td>...</td>
</tr>
<tr>
<td>Left to right hemispheric asymmetry of atrophy</td>
<td>37</td>
<td>...</td>
</tr>
<tr>
<td>Left to right asymmetry or anterior greater than posterior gradient of atrophy</td>
<td>59</td>
<td>...</td>
</tr>
</tbody>
</table>
metrical hippocampal atrophy (any severity) was 71% sensitive for AD and 70% and 47% specific for AD when compared with controls and patients with FTLD, respectively. Only posterior greater than anterior gradient of atrophy was highly specific for AD when compared with other pathologies, and sensitivity was moderate. Combining bilateral symmetrical hippocampal atrophy and posterior greater than anterior gradient of atrophy gave 87% sensitivity for AD, with moderate specificity when compared with other pathologies.

All 17 patients with FTLD pathology were distinguished from controls after MRI assessment, and a mean (SD) of 61% (14.8%) were correctly diagnosed as having FTLD pathology. Of those misdiagnosed, 85% were thought to be AD. Atrophy (any severity) of the ATL or amygdala, LTG or fusiform gyrus, and parahippocampal gyrus had high sensitivity (≥90%) for FTLD; specificity was moderate but decreased to 32%, 35%, and 39%, respectively, when FTLD was compared with AD (Table 2). Considering moderate to severe atrophy of these regions increased specificity for FTLD compared with AD to more than 75%, but sensitivity was reduced. Frontal lobe atrophy had moderately good sensitivity (76%) and specificity for FTLD when compared with all other cases and AD cases (74% and 60%, respectively). Atrophy patterns specific for FTLD when compared with other pathologies were anterior greater than posterior gradient of atrophy and asymmetry of atrophy; however, sensitivity was less than 40%. Combining anterior greater than posterior gradient of atrophy and asymmetry of atrophy resulted in a higher sensitivity of 59% for FTLD, while specificity remained high at 84% when AD and FTLD were compared.

Table 3 lists the results of multiple logistic regression analysis performed on AD vs controls, AD vs all other cases, and AD vs FTLD cases. When comparing atrophy patterns in AD and controls, 11 assessments of AD cases rated as having asymmetry of atrophy, a feature that is 100% specific for AD, were excluded from the analysis. Subsequent analysis identified frontal lobe, ATL or amygdala, and hippocampal atrophy as statistically significant independent predictors of AD compared with controls. In distinguishing AD from all other cases, statistically significant atrophy patterns were posterior greater than anterior gradient of atrophy and hippocampal atrophy, which multiplied the odds of having AD by factors of around 5.0 and 1.5, respectively. Isolated unilateral parahippocampal gyrus atrophy was associated with a reduced risk of having AD compared with other diagnoses. In assessing atrophy patterns distinguishing AD from FTLD, posterior greater than anterior gradient of atrophy was statistically significant, with the odds of having AD multiplied 10-fold. In addition, anterior greater than posterior gradient of atrophy and parahippocampal gyrus atrophy were statistically significant independent predictors of a diagnosis of FTLD.

This study investigated the capacity of simple visual inspection of MRIs to distinguish diseases causing young-onset dementia in patients with a range of pathologically confirmed dementing diseases. We demonstrated that visual assessment of MRIs is a sensitive method for identifying dementia and that there are several global and regional atrophy patterns more commonly associated with a pathologically confirmed diagnosis of AD or FTLD.

Hippocampal atrophy of any severity was 92% sensitive and 62% specific for AD compared with controls. Our results are similar to those found in a study of clinically diagnosed AD and controls that rated medial temporal lobe atrophy and found 81% sensitivity and 67% specificity for AD. However, our study found that specificity of hippocampal atrophy for AD was poor when AD and FTLD were compared, and it was a nonsignificant factor in multiple logistic regression analysis. Although moderate to severe atrophy of the hippocampus can discriminate AD from controls with 98% specificity, discrimination from other pathologies remained poor, and sensitivity was reduced. This is in accord with studies that demonstrated hippocampal atrophy in FTLD in vascular dementia and normal aging. The pattern of atrophy within the hippocampus is important, with bilateral symmetrical hippocampal atrophy occurring in AD, in contrast to asymmetrical atrophy that occurs in FTLD. In agreement with this, bilateral symmetrical hippocampal atrophy had the highest sensitivity and specificity for AD when compared with FTLD. Atrophy of the parietal cortex has also been associated with AD and although sensitivity was moderate in our study, posterior greater than anterior gradient of atrophy was highly specific for AD, especially when com-
pared with FTLD. Combining bilateral symmetrical hippocampal atrophy and posterior greater than anterior gradient of atrophy may ensure a higher specificity for discriminating AD from FTLD and other pathologies than that achieved by hippocampal atrophy alone.

Although several pathologic substrates underlie a clinical diagnosis of FTLD, this study based diagnosis on 3 established clinical subtypes (FLD, primary progressive nonfluent aphasia, and semantic dementia)\(^\text{14}\) that appear to be associated with different atrophy patterns involving the frontal or temporal lobes, often with left to right asymmetry.\(^\text{11,13,14,30}\) In our study, the ATL or amygdala, LTG or fusiform gyrus, and parahippocampal gyrus were rated atrophic in 90% or more of the patients with FTLD, and multiple logistic regression analysis identified atrophy of the parahippocampal gyrus as a significant predictor of FTLD. Previous studies have identified the amygdala,\(^\text{11,13}\) LTG or fusiform gyrus,\(^\text{3,30}\) and parahippocampal gyrus\(^\text{13,36}\) as commonly affected in clinically diagnosed FTLD. Chan et al\(^\text{13}\) found fusiform gyrus atrophy to be highly specific for semantic dementia compared with AD; in contrast, our study found the specificity for FTLD (FLD, primary progressive nonfluent aphasia, or semantic dementia) of atrophy in these regions was low to moderate, highlighting the overlap of atrophy patterns in pathologically confirmed AD and FTLD. Early changes in the parahippocampal gyrus (particularly the entorhinal cortex) have been associated with AD\(^\text{13,34}\), reports suggest that atrophy is more severe in FTLD, particularly on the left.\(^\text{11,13,29,32}\) A study\(^\text{16}\) comparing semantic dementia, FLD, and AD found that moderate to severe parahippocampal gyrus atrophy was highly sensitive for semantic dementia and found in only 17% of the subjects with AD. In our study, moderate to severe atrophy of the ATL or amygdala, LTG or fusiform gyrus, and parahippocampal gyrus each increased specificity for FTLD to more than 75%, with moderate sensitivity, which may reflect the fact that our FTLD group consisted of more FLD and primary progressive nonfluent aphasia than semantic dementia clinical variants, as Galton et al\(^\text{16}\) found that a much smaller proportion of subjects with FLD than those with semantic dementia had moderate to severe ATL or LTG atrophy.

Compared with other pathologies, anterior greater than posterior gradient of atrophy and asymmetry of atrophy were highly specific for FTLD but of low sensitivity, which may reflect the difficulty of assessing global atrophy patterns in these patients, some of whom may have only mild atrophy. A recent study\(^\text{22}\) comparing clinically diagnosed FTLD, AD, and vascular dementia found a similarly high specificity and low sensitivity (100% and 38%, respectively) for FTLD of asymmetrical atrophy. Rating of severe frontal atrophy in these cases demonstrated that it was 93% specific and 52% sensitive. The study\(^\text{20}\) also showed that combining asymmetry and frontal atrophy ratings increased sensitivity to 71%, with specificity remaining high at 93%, a result similar to ours when combining anterior greater than posterior gradient of atrophy and asymmetry of atrophy. Another study\(^\text{16}\) found that a discriminant function that includes values of frontal or temporal asymmetry could distinguish FTLD from AD with 90% sensitivity and 93% specificity. Anterior greater than posterior gradient of atrophy, asymmetry of global atrophy, and moderate to severe ATL or amygdala atrophy may be additive in their detection of FTLD. Indeed, the highest sensitivity among the 3 observers for the detection of FTLD was 76%, with a similarly high specificity.

Findings from our study can also be compared with those obtained from region of interest measurements, a more precise quantification of atrophy. Several studies\(^\text{29,34,36}\) investigating hippocampal volume measurements to classify subjects with clinically diagnosed AD from controls found that sensitivity ranged from 75% to 80%, with specificity ranging from 76% to 90%. In contrast, our visual inspection study found slightly lower sensitivity but higher specificity. Investigating discrimination of subjects with clinically diagnosed FTLD from controls, one study\(^\text{20}\) found that volumetric measures of entorhinal cortex distinguished only 50% of subjects, for a specificity of 90%. This contrasts with the 92% sensitivity and 61% specificity found in our study of visual ratings of parahippocampal gyrus to discriminate FTLD from other dementias and controls (arguably a more difficult distinction). Few volumetric studies have attempted to use measures to discriminate between different diseases.\(^\text{13,34,36}\) One study looked at the benefit of using volumetric measures or visual ratings of the medial temporal lobe in addition to MMSE scores to discriminate AD from other dementias.\(^\text{14}\) Volumetric measures yielded no diagnostic gain over the MMSE scores (68% sensitivity and 53% specificity), whereas visual ratings did (78% sensitivity and 64% specificity).\(^\text{14}\) Although quantitative techniques may be more precise in their measurement of atrophy, unlike visual assessment, they do not consider the relative state of atrophy within the whole brain and are more labor and equipment intensive. Visual inspection appears to perform well in comparison, especially given that the technique is more applicable in a clinical setting.

The mix of patients in this study, young age at onset, and inclusion of atypical presentations (postmortem or biopsy is more likely to be performed when a clinical diagnosis is uncertain) enabled visual inspection of MRIs to be tested in a demanding and clinically realistic manner. Furthermore, although mixed pathology is common and it is important for the primary pathology to be identified, radiological diagnosis may have been more challenging in these cases. Although the patients in our study had imaging performed as part of a diagnostic workup, the mean duration from earliest symptom to MRI was around 4 years. Although early symptoms may be subjective, subtle, and only elicited on direct questioning of the patient and relatives, MRI observations would be most useful at this time. The inclusion of patients at a more advanced stage of disease may reduce the relevance of our findings to the earliest detection of dementias but reflects clinical practice, in which there may be a considerable delay between first symptoms and brain imaging and finally to diagnosis. A previous study has shown a mean time from symptoms to diagnosis of around 4 years.\(^\text{37}\) Although the difference in MMSE scores between the AD and FTLD groups suggests that patients with AD were more severely affected, this may instead reflect the limitations of assessing disease severity in FTLD using MMSE scores.\(^\text{14}\)
Accepted for Publication: March 24, 2005.

Correspondence: Valerie M. Anderson, BSc, Dementia Research Centre, Institute of Neurology, Queen Square, London WC1N 3BG, England (vanderson@dementia.ion.ucl.ac.uk).


Funding/Support: This study was funded by EU contract QLK3-CT-2001-02362/LSHM-CT-2003-503330 (administered by VERUM Foundation, Munich, Germany) from Early Diagnosis of Alzheimer’s Disease and Related Dementias/Abnormal Proteins in the Pathogenesis of Neurodegenerative Disorders (DIADEMA/APOPIS); the Alzheimer’s Disease Society, London; Alzheimer’s Research Trust, Cambridge, England; and Medical Research Council, London. Dr Fox holds a Medical Research Council Senior Clinical Fellowship.

Additional Information: Dr Likeman and Ms Anderson contributed equally to this study.

Acknowledgment: We thank the patients and their families for consenting to biopsy and brain donation. We acknowledge the following neuropathology departments for performing postmortem and histologic examinations: Institute of Psychiatry, London; National Hospital for Neurology and Neurosurgery; Queen Square Brain Bank for Neurological Disorders, London; Addenbrooke’s Hospital, Cambridge; and Radcliffe Infirmary, Oxford, England.

REFERENCES


©2005 American Medical Association. All rights reserved.