Differential diagnosis of Alzheimer’s disease and vascular dementia using visual rating scales

Stamo Manouvelou¹, Vasileios Koutoulidis¹, Maria Tolia², Athanasios Gouliamos³, Georgios Anyfantakis¹, Lia-Angela Moulopoulos¹, Sokratis G. Papageorgiou¹

¹1st Department of Radiology, University of Athens Medical School, Aretaieion University Hospital, Athens, Greece
²Department of Radiotherapy/Radiation Oncology, Faculty of Medicine, School of Health Sciences, University of Thessaly, University Hospital, Larisa, Greece
³2nd Department of Neurology, Attikon University Hospital, University of Athens Medical School, Athens, Greece

Abstract

Purpose: To determine the differential diagnosis of Alzheimer’s disease (AD) and vascular dementia (VaD) using visual rating scales.

Material and Methods: 119 brain Magnetic Resonance Imaging (MRI) examinations of patients diagnosed either with AD (n=85) or VaD (n=34) were assessed. Double blinded visual evaluation was performed by two neuroradiologists. Final clinical diagnosis was set by a behavioural neurologist. The following rating scales were valued: Pasquier rating scale (GCA), Fazekas Scale assessing both periventricular (PV) and white matter (WM) lesions. Posterior Cortical Atrophy (PCA) and scales regarding specific cortical regions: dorso-frontal (DF), orbito-frontal (OF), anterior-cingulate (AC), basal ganglia (BG), anterior-temporal (AT), insula, lateral-temporal (LT), entorhinal (ERC), perirhinal (PRC), anterior-fusiform cortex (AFC), anterior-hippocampus (AHIP) and posterior-hippocampus (PHIP). Both Left (L) and Right (R) hemispheres were examined.

Results: The indicators with the highest value of area under the curve (AUC) were Fazekas-WM (AUC: 0.906), Fazekas-PV (AUC: 0.894), R-ERC (AUC: 0.858) and L-ERC (AUC: 0.820). Best sensitivity for distinguishing VaD from AD was achieved by PCA (91%), R-ERC (91%), Fazekas-WM (89%) and L-ERC (87%). Highest specificity was achieved by Fazekas-WM (97%), R-PHIP (91%), L-PHIP (91%), R-AC

Corresponding Author: Stamo Manouvelou, 1st Department of Radiology, University of Athens Medical School, Aretaieion University Hospital, 76 Vassilisis Sofias, 115 28 Athens, Greece, Email: smanouvelou@yahoo.gr
Guarantor: Prof. Sokratis G. Papageorgiou, 2nd Department of Neurology, Attikon University Hospital, University of Athens Medical School, 1 Rimini, 12462, Athens, Greece, Email address: sokpapa@med.uoa.gr
Dementia/vascular; Alzheimer Disease; MR imaging/diagnosis; Brain disorders

Introduction
The interaction between white matter hyperintensities (microangiopathy) detected on Magnetic Resonance Imaging (MRI), cerebrovascular risk factors and cognitive decline in the form of either Alzheimer’s disease (AD) or vascular dementia (VaD) is complex. Cognitive decline and the presence of white matter hyperintensities is well established by many cross-sectional and longitudinal studies [1-4]. Furthermore, recent studies have detected not only associations between vascular risk factors, such as hypertension, diabetes mellitus, hyperlipidaemia and AD, but also association between cerebrovascular disease and AD. In fact, common aetiologic or reciprocally synergistic pathophysiological mechanisms are shared between AD and VaD [5-12].

Differential diagnosis of VaD from AD is a difficult task since MRI scans may show white matter hyperintensities in both entities. Furthermore, white matter hyperintensities are often age related, which is the case for both VaD and AD [13]. A combination of clinical and neuroimaging evaluations can maximise the accuracy of the diagnostic process. Structural imaging based on MRI is an integral part of the clinical assessment of patients with dementia. The aim of our study is to identify an imaging pattern that may help clinicians in the differential diagnosis of AD and VaD.

Material and Methods
In our retrospective clinical study, brain MRI examinations of 119 patients diagnosed with AD (n=85, 71.4%) and VaD (n=34, 28.6%) were evaluated. The examinations were performed for the diagnostic evaluation of patients hospitalised at the Cognitive Disorder/Dementia Unit Clinic of our department of Neurology during the period from 2004 to 2012. The study was approved by the Ethics Committee of our hospital. The examinations were performed with 1.5 Tesla MRI scanners using standard brain MRI protocol. The essential sequences were coronal 3D T1-weighted gradient echo (1 mm isotropic voxels), transverse T2WI TSE/FSE (3-5 mm slices), transverse FLAIR TSE/FSE (3-5 mm slices) and transverse T2* gradient echo (3-5 mm slices).

The images were retrospectively evaluated using mainly FLAIR and T2WI axial imaging for the evaluation of white matter hyperintensities and coronal/axial FLAIR and T1WI for the evaluation of cerebral atrophy. Double random blinded visual evaluation in order to assure intra-rater and inter-rater agreement was performed by two trained neuroradiologists (S.M. and A.G.). The time interval between the two evaluations in order to assess intra-observer variability was 6 months for both neuroradiologists. The clinical diagnosis was set by a behavioural neurologist with specific expertise in dementia (S.G.P.), who was not aware of the MRI findings. The possible diagnoses were only AD and VaD. Patients with mixed pattern of AD and VaD were excluded from the study, as were patients with cerebrovascular large vessel. The visual rating model used was adjusted for differences in age, gender, education, whole brain volume and Mini-Mental State Examination (MMSE). Disease duration and MMSE score were also evaluated.

The following rating scales were evaluated:
- Pasquier rating scale for Global Cortical Atrophy (GCA) evaluation, which represents strong correlation with the severity of AD and Mild Cognitive Impairment (MCI). Global volume loss without focal lobar atrophy is a common and non-specific finding on structural MRI studies in normal ageing and dementia. Visual rating of cortical atrophy can be easily performed using a four-point (0-3) Pasquier scale for the detection of the presence and extent of atrophy.
- Fazekas-PV rating scale for the detection of periventricular hyperintensities. This scale is based on the assessment of the intensity and extent of hyperintensities in the periventricular region. It is a commonly used and validated rating scale for the detection of white matter hyperintensities.

Key words
- Dementia/vascular
- Alzheimer Disease
- MR imaging/diagnosis
- Brain disorders
Differential diagnosis of Alzheimer’s disease and vascular dementia using visual rating scales, p. 2-9

Based on the width of the sulci and the volume of the gyri [14]. Visual scales regarding specific brain regions: dorso-frontal (DF), orbito-frontal (OF), anterior cingulated (AC), anterior or temporal (AT), insula, lateral temporal (LT), entorhinal (ERC), perirhinal (PRC) and anterior-fusiform cortex (AF) posterior hippocampus (PHIP) and anterior hippocampus (AHIP), a rating scale proposed by Davies et al [15]. Both hemispheres were separately evaluated.

Fazekas White Matter Changes Scale: as an easily applicable and highly reproducible four-point rating scale. It is widely used in clinical practice and corresponds well with more detailed rating scales and histopathology, representing an important marker for the evaluation of vascular dementia [16]. Hyperintensities within the cerebral white matter on T2-weighted or FLAIR imaging and less prominently on T1-weighted imaging are more likely to be vascular in origin. However, even in cases of extensive white matter hyperintensities, the existence of mixed pathology should be considered, although it may be difficult to confirm or refute. White matter hyperintensities can be divided in periventricular (PV) and white matter (WM) lesions, the first regarding lesions in continuity with the margins of each lateral ventricle, while the second regarding lesions separate from the ventricles [16].

Koedam rating scale for Posterior Cortical Atrophy evaluation involving the parietal/occipital cortex is usually the result of underlying AD pathology. Visual rating of posterior atrophy in combination with MTL atrophy rating has been reported to help discriminating AD from Fronto temporal dementia with a sensitivity of 73% and a specificity of 87% [17].

Krainik rating scale is an anteroposterior gradient of atrophy, using a 5 point-scale (-2 to 3) for evaluation of a sagittal midbrain slice, which can help distinguish different types of dementia based on the anteroposterior gradient of cortical atrophy. Anterior atrophy prevalence was observed in Frontotemporal dementia, whereas Alzheimer’s disease was present at a more posterior gradient [18].

Table 1. Demographics

<table>
<thead>
<tr>
<th></th>
<th>VaD</th>
<th>AD</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>67.5 ± 8.8</td>
<td>67.6 ± 9.3</td>
<td>0.959</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>20 (58.8%) /14 (41.2%)</td>
<td>50 (58.8%) /35 (41.2%)</td>
<td>1.000</td>
</tr>
<tr>
<td>Education (years)</td>
<td>9.10 ± 4.87</td>
<td>10.31 ± 4.44</td>
<td>0.102</td>
</tr>
<tr>
<td>Disease duration [median years (range)]</td>
<td>2.0 (0-17)</td>
<td>2.4 (0-7)</td>
<td>0.603</td>
</tr>
<tr>
<td>MMSE</td>
<td>19.77 ± 5.32</td>
<td>18.67 ± 5.79</td>
<td>0.367</td>
</tr>
<tr>
<td>Predominant hand (right/left)</td>
<td>30 (91.2%)/3 (8.8%)</td>
<td>77 (90.5%)/8 (9.5%)</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Fig. 1. ROC analysis (VaD vs AD) AUC>0.7.

Table 1. Demographics

<table>
<thead>
<tr>
<th></th>
<th>VaD</th>
<th>AD</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>67.5 ± 8.8</td>
<td>67.6 ± 9.3</td>
<td>0.959</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>20 (58.8%) /14 (41.2%)</td>
<td>50 (58.8%) /35 (41.2%)</td>
<td>1.000</td>
</tr>
<tr>
<td>Education (years)</td>
<td>9.10 ± 4.87</td>
<td>10.31 ± 4.44</td>
<td>0.102</td>
</tr>
<tr>
<td>Disease duration [median years (range)]</td>
<td>2.0 (0-17)</td>
<td>2.4 (0-7)</td>
<td>0.603</td>
</tr>
<tr>
<td>MMSE</td>
<td>19.77 ± 5.32</td>
<td>18.67 ± 5.79</td>
<td>0.367</td>
</tr>
<tr>
<td>Predominant hand (right/left)</td>
<td>30 (91.2%)/3 (8.8%)</td>
<td>77 (90.5%)/8 (9.5%)</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Fig. 1. ROC analysis (VaD vs AD) AUC>0.7.

and highly reproducible four-point rating scale. It is widely used in clinical practice and corresponds well with more detailed rating scales and histopathology, representing an important marker for the evaluation of vascular dementia [16]. Hyperintensities within the cerebral white matter on T2-weighted or FLAIR imaging and less prominently on T1-weighted imaging are more likely to be vascular in origin. However, even in cases of extensive white matter hyperintensities, the existence of mixed pathology should be considered, although it may be difficult to confirm or refute. White matter hyperintensities can be divided in periventricular (PV) and white matter (WM) lesions, the first regarding lesions in continuity with the margins of each lateral ventricle, while the second regarding lesions separate from the ventricles [16].

Koedam rating scale for Posterior Cortical Atrophy evaluation involving the parietal/occipital cortex is usually the result of underlying AD pathology. Visual rating of posterior atrophy in combination with MTL atrophy rating has been reported to help discriminating AD from Fronto temporal dementia with a sensitivity of 73% and a specificity of 87% [17].

Krainik rating scale is an anteroposterior gradient of atrophy, using a 5 point-scale (-2 to 3) for evaluation of a sagittal midbrain slice, which can help distinguish different types of dementia based on the anteroposterior gradient of cortical atrophy. Anterior atrophy prevalence was observed in Frontotemporal dementia, whereas Alzheimer’s disease was present at a more posterior gradient [18].

Statistical analysis

Data are expressed as mean ± standard deviation (S.D.) or median (in case of skewed distribution) for continuous variables and as percentages for categorical variables. The Kolmogorov-Smirnov test was utilised for normality analysis of the parameters.

In order to quantify inter- and intra-observer agreement, we calculated linear weighted- kappa’s between each pair of observers and between the first and second session of the first ratings separately. For weighted-kappa values, degree of agreement was defined according to Landis and Koch.

The main outcome of the study was the Area under the curve (AUC) for the AD status compared to the VaD status. Sensitivity, specificity, positive predictive values and negative predictive values were obtained from the visual rating cut-offs. Receiver operator characteristic
(ROC) curves were constructed to examine sensitivity and specificity for all combinations of rating cut-off scores.

Logistic-regression analysis was used to analyse relationships between scores on each visual rating scale, with ordinal or categorical form and variable outcome (AD vs. VaD). In the first model all visual rating scales with AUC>0.7 were entered simultaneously in the model. In the second model all visual rating scales with AUC>0.7 were entered in the model and stepwise elimination (Wald method) was used to reach the final model. The results presented are logistic regression coefficients with 95% CI, OR (95% CI).

All tests are two-sided and statistical significance was set at p<0.05. All analyses were carried out using the statistical package SPSS v 19.0 (IBM SPSS Statistics for Windows, Version 19.0 Armonk, NY: IBM Corp).

Results
No statistical differences between the demographic characteristics of two sub-populations was observed. The demographics of the two populations are presented in Table 1.

Intraobserver reliability
The linear weighted- kappa’s ranged from 0.756 (95% CI: 0.69- 0.83) for BG to 0.979 (95% CI: 0.96-0.99) for Fazekas-PV, which represents the highest agreement between the first and the second evaluation of the main rater.

Interobserver reliability
The linear weighted kappa’s ranged from 0.772 (95% CI: 0.71-0.83) for R-AC to 0.958 (95% CI: 0.94-0.98) for Fazekas-WM, which represents the greatest agreement between the first evaluation of the first rater and the first evaluation of the second rater. Ratings with kappa value >0.9 were GCA, Table 2. ROC analysis (AD vs VD) reliability Indicators

<table>
<thead>
<tr>
<th>Cut-off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fazekas-PV</td>
<td>2.5</td>
<td>89%</td>
<td>82%</td>
<td>92%</td>
</tr>
<tr>
<td>Fazekas-WM</td>
<td>1.5</td>
<td>70%</td>
<td>97%</td>
<td>98%</td>
</tr>
<tr>
<td>RPHIP</td>
<td>2.5</td>
<td>57%</td>
<td>91%</td>
<td>94%</td>
</tr>
<tr>
<td>LPHIP</td>
<td>2.5</td>
<td>47%</td>
<td>91%</td>
<td>93%</td>
</tr>
<tr>
<td>PCA</td>
<td>1.5</td>
<td>91%</td>
<td>53%</td>
<td>83%</td>
</tr>
<tr>
<td>R-AC</td>
<td>1.5</td>
<td>34%</td>
<td>91%</td>
<td>91%</td>
</tr>
<tr>
<td>L-AC</td>
<td>1.5</td>
<td>33%</td>
<td>91%</td>
<td>91%</td>
</tr>
<tr>
<td>R-LT</td>
<td>1.5</td>
<td>72%</td>
<td>59%</td>
<td>59%</td>
</tr>
<tr>
<td>L-LT</td>
<td>1.5</td>
<td>68%</td>
<td>62%</td>
<td>62%</td>
</tr>
<tr>
<td>R-AHIP</td>
<td>2.5</td>
<td>64%</td>
<td>73%</td>
<td>86%</td>
</tr>
<tr>
<td>L-AHIP</td>
<td>2.5</td>
<td>51%</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>R-ERC</td>
<td>1.5</td>
<td>91%</td>
<td>62%</td>
<td>87%</td>
</tr>
<tr>
<td>L-ERC</td>
<td>1.5</td>
<td>87%</td>
<td>62%</td>
<td>85%</td>
</tr>
<tr>
<td>R-PRC</td>
<td>1.5</td>
<td>68%</td>
<td>73%</td>
<td>87%</td>
</tr>
<tr>
<td>L-PRC</td>
<td>1.5</td>
<td>62%</td>
<td>73%</td>
<td>85%</td>
</tr>
<tr>
<td>R-AF</td>
<td>0.5</td>
<td>80%</td>
<td>41%</td>
<td>41%</td>
</tr>
<tr>
<td>L-AF</td>
<td>0.5</td>
<td>79%</td>
<td>41%</td>
<td>41%</td>
</tr>
</tbody>
</table>

Table 3. Multiple logistic regression model-Forward Wald selection (categorical indicators)

<table>
<thead>
<tr>
<th>Reference category</th>
<th>Odds Ratio</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fazekas PV</td>
<td>2.5</td>
<td>118.25</td>
<td>14.47</td>
</tr>
<tr>
<td>R-ERC</td>
<td>1.5</td>
<td>56.87</td>
<td>6.54</td>
</tr>
</tbody>
</table>
Fazekas-PV and Fazekas-WM. Ratings with kappa value >0.8 were R-OF, L-OF and R-AC.

Fig. 1 and Table 2 present the ROC analysis, sensitivity, specificity and the cut-off points of the indicators for distinguishing VaD from AD. We observe that the indicator with the highest value of AUC was Fazekas-WM (AUC: 0.906) followed by Fazekas-PV (AUC: 0.894), R-ERC (AUC: 0.858) and L-ERC (AUC: 0.820).

In relation to the sensitivity of the indicators for distinguishing VaD from AD the best value was presented by PCA (91%) and R-ERC (91%) followed by Fazekas-PV (89%), L-ERC (87%) and R-AF (80%). The highest value of specificity in the discrimination of VaD from AD was presented by Fazekas-WM (97%) followed by R-PHIP (91%); L-PHIP (91%) and R-AC (91%). Finally, the best combination of sensitivity and specificity for distinguishing VaD from AD was presented by Fazekas-PV (89% -82%) followed by Fazekas-WM (70%-97%) and R-ERC (91%-62%). The best combination of PPV and NPV was observed by Fazekas WM (92%-76%), followed by R-ERC (87%-72%) and Fazekas-PV (98%-56%).

Table 3 presents the analysis of multiple logistic regression for the variable division of VaD from AD (appearance of AD compared to VaD). Using a multiple linear regression model with the method of forward selection Wald, including only the indicators with AUC>0.7 and trying to detect the independent influence of the strongest predictors on the occurrence of variable AD relative to VaD, we observed that the indices Fazekas-PV (p<0.0005) and R-ERC (p=0.004) have a statistically significant effect on the variable occurrence AD compared to VaD.

Table 4 presents the analysis of multiple logistic regression for the variable division of VaD from AD (AD appearance relative to VaD). Using multiple linear regression model with the method of introducing all variables with AUC>0.7 (Multiple logistic regression-enter method) and trying to depict the independent effect of the indicators of these tools on the variable appearance AD compared to VaD, we observe that the indices Fazekas WM (p=0.008) and ERC-R (p=0.004) have a statistically significant effect on the variable occurrence AD relative to VaD.

The interpretation of the indicators is the following:

a) Increasing Fazekas PV index by 1 unit reduces the probability of having AD vs VaD by 91%, and b) Increasing ERC-R index by 1 unit brings about 16 times probability of having AD vs VaD.

Discussion
In our study Fazekas-WM and Fazekas-PV helped distinguish better VaD from AD. Fazekas-WM, entorhinal cortex and hippocampal atrophy helped in the differential diagnosis of Alzheimer’s disease, in accordance to studies presented in the literature [14-26]. By evaluating standard clinical FLAIR and T2WI brain MRI studies and focusing on white matter hyperintensities in addition to other vascular findings such as lacunar infarcts, microbleeds, and focal encephalomalacia, diagnosis of vascular dementia becomes more probable. Similarly, by evaluating FLAIR and T1WI sequences with focus on atrophy of the hippocampal and en-
torhinal region, diagnosis of AD becomes likely. Our results indicate that Fazekas PV and Fazekas WM were more likely to be associated with a diagnosis of VaD than AD. Additionally, hippocampal and entorhinal atrophy indicators are associated with the diagnosis of AD rather than VaD, with very good combination of PPV NPV.

The presence of anterior cingulate atrophy in our findings, which achieved good specificity (91%), may be due to a misdiagnosis of patients with Frontotemporal lobe degeneration (FTD), in particular behavioural variant (bvFTD) with AD. Clinically, FTD is usually clearly distinguishable from the typical amnestic AD presentation. However, there can be a grey area with some FTD patients having prominent impairment of episodic memory. In fact, for decades the diagnostic criteria of the various clinical types of FTD were those of excluding cases that followed a pattern of symptomatology characteristic of AD and VaD, but recent accumulated evidence indicates that episodic memory impairment, hallmark of AD, could be the initial clinical expression of bvFTD [27]. Cingulate atrophy is present in both AD and bvFTD, with anterior cingulate atrophy involved in bvFTD and posterior cingulate atrophy in AD [28-29].

There are few limitations to our study. Firstly, its retrospective nature and secondly the fact that the diagnosis of dementia was made solely on the basis of the clinical picture of the patient, by a behavioural neurologist with no imaging modalities available. Furthermore, the diagnosis has not been confirmed by autopsy, thus some patients could have been misdiagnosed. Nevertheless, recent studies have histopathologically validated the accuracy of visual scales mainly used in our study [30].

It must be noted that our patients were evaluated in a memory disorder clinic where the frequency of AD was high, and in this setting our findings are most relevant. It should be kept in mind that the number of patients with VaD was relatively small. Furthermore, the disadvantages of the rating approach are obvious, since it will not give precise volumes. Rating is prone to sampling error since only
one view of a structure is assessed. Furthermore, rating lacks objectivity, although it can be improved with rigorous blinding. In our study in fact the good intra-rater and inter-rater agreement rates of the rating scales must be noted, supporting their validity as a diagnostic tool.

It must be acknowledged that the proposed methodology can be time consuming due the fact that it involves many rating scales and may not be easily applicable in everyday practice. Nevertheless, an effort to use Fazekas scale and entorhinal cortex and hippocampal atrophy evaluation in the diagnosis of dementia is valid to be made.

MRI measures of atrophy and white matter hyperintensities reflect neuronal damage which is directly responsible for patients’ clinical status. When compared to other imaging markers (and other biomarkers) cerebral atrophy has an advantage, which is its strong correlation with cognitive decline [4, 10, 31, 32]. Visual rating scales in the form of focal atrophy or hyperintensities evaluation may offer information leading towards the diagnosis of a specific form of dementia (Figs. 2-4). These rating scales are inexpensive, fast and easy to apply. Accurate and timely diagnosis is increasingly important to guide management and to provide appropriate information and support patients with dementia. In the last years, neuroimaging has contributed to the phenoypic characterisation of such patients. Structural imaging in cognitive cases can provide accessible, clinically useful information. It is the primary neuroimaging technique of choice in clinical practice to support the clinical diagnosis of dementia. Visual rating scales can become a useful cost-effective tool in the differential diagnosis of patients with AD and VaD and can be used as an alternate to highly advanced but expensive and time consuming current technical methods of quantification.

Acknowledgment
Stamo Manouvelou was awarded a Fellowship of excellence for postgraduate studies in Greece-Siemens program.

Conflict of interest
The authors declared no conflicts of interest.

REFERENCES


