

# Grading white matter lesions on CT and MRI: a simple scale

J C van Swieten, A Hijdra, P J Koudstaal, J van Gijn

## Abstract

**We developed and tested a simple three-point scale for grading white matter lesions in anterior and posterior regions of the brain. Twenty four CT scans and 24 MRI scans were separately judged by 11 and five observers, respectively, on the presence and severity of white matter lesions. The observers were radiologists and neurologists. For CT scans, these periventricular changes were graded according to their extent as absent, or partly involving the white matter, or extending up to the sub-cortical region. The MRI lesions were graded as no lesion or only a single one, multiple focal lesions, and multiple confluent lesions. The pairwise agreements of all possible combinations of observers for each scan were corrected for chance (kappa statistics; maximal agreement 1.0). The weighted kappa value, for anterior and posterior regions combined, was 0.63 for CT scans, and 0.78 for MRI scans. This three-point scale for two separate regions seems suitable as a basis for cross-sectional or longitudinal studies of large series of patients.**

Periventricular leukoencephalopathy is a more or less symmetrically diminished density of the white matter on CT scans of elderly individuals.<sup>1-5</sup> The frontal and the parieto-occipital white matter are most often affected, although not always to equal degrees,<sup>6</sup> frequently in association with ventricular enlargement. Magnetic resonance imaging (MRI) has proved more sensitive in detecting white matter lesions than CT.<sup>7-9</sup> The hypodense lesions on CT correspond with areas of increased signal intensity in T2 weighted MRI. They are most often found in patients with stroke, hypertension or dementia.<sup>10-16</sup> Neuropathological studies have shown that these lesions on CT or MRI correspond with demyelination, astrocytic gliosis, arteriolosclerosis, or dilated perivascular spaces, and are often associated with lacunar infarcts.<sup>17,18</sup>

Further studies of the incidence, causes and prognosis of periventricular white matter lesions are impeded by the lack of reliable criteria for the presence and severity of these lesions. The presence of white matter lucencies on CT has in some cases been quantified in Hounsfield units,<sup>2,3</sup> but in most

previous studies a subjective assessment was made of the extent or degree of attenuation of these lesions at one site, compared with the surrounding brain.<sup>10,12,17,19,20</sup> Other CT studies have distinguished different types according to the anatomical distribution of white matter lesions, but did not take account of the severity.<sup>1,3</sup>

In MRI studies, a variety of scales has been used for measuring the severity of white matter lesions. In most cases two categories of lesions were distinguished that often occur together, that is, periventricular lesions and focal lesions.<sup>14-16,21-26</sup> Sometimes these lesions have been separately graded<sup>22,24</sup> and elsewhere they were taken together.<sup>21,23,25</sup> The extent of the periventricular lesions in the white matter may range from narrow rims around the frontal horns to thick bands extending deep in the white matter, and focal lesions may be single, multiple or confluent.

Whether these white matter lesions imply functional impairment is not entirely clear. Studies about cognitive functions in patients with white matter lesions on CT or MRI have given conflicting results.<sup>19,20,24-28</sup> Although this may have been caused by differences in patient selection, another and perhaps neglected factor may be the severity of the lesions in the frontal or parieto-occipital white matter. In future studies in which periventricular leukoencephalopathy is related to function or to future events, white matter lesions on CT or MRI should be graded according to both the anatomical pattern and the severity.

An ideal rating scale for white matter lesions on CT or MRI should meet the following requirements. First, it should include the anatomical distribution as well as the severity of white matter lesions, with clear definitions for the different categories. Secondly, the scale should be simple. Finally, the reliability of the scale should be assessed in an inter-observer study. One such study has been reported, but without mention of the type of scale or the actual measure of agreement.<sup>29</sup> In this study, we have developed a scale for CT and another for MRI, consisting of similar categories for anterior and posterior regions, but with different definitions for the kind of lesions on CT or MRI. We determined the extent of agreement among different observers by means of kappa statistics, which accounts for chance agreement.

## Methods

A method for systematic evaluation of the

Department of  
Neurology, University  
Hospital, Utrecht  
J C van Swieten  
J van Gijn

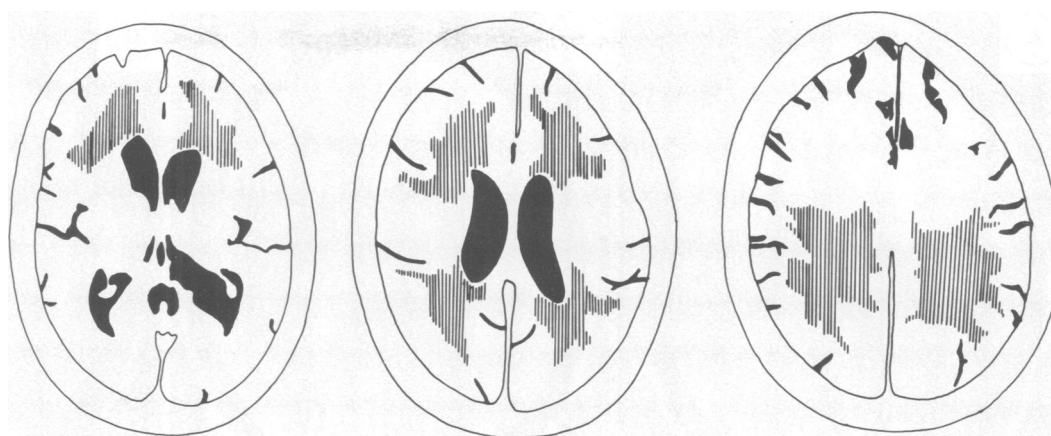
Department of  
Neurology, Academic  
Medical Centre,  
Amsterdam  
A Hijdra

Department of  
Neurology, University  
Hospital Dijkzigt,  
Rotterdam, The  
Netherlands  
P J Koudstaal

Correspondence to:  
Dr van Swieten,  
Department of Neurology,  
University Hospital,  
Heidelberglaan 100,  
3584 CX Utrecht,  
The Netherlands.

Received 31 May 1989  
and in revised form  
9 February 1990.  
Accepted 3 April 1990

**Figure 1** The severity of white matter lesions is graded as 0, 1 or 2 for the anterior white matter in the slice through the choroid plexus (left) and the slice through the cella media (centre), and for the posterior white matter in the slice through the cella media (middle) and the slice through the centrum semiovale (right), for instance,  $A_0$ ,  $P_1$ , or  $A_2$ ,  $P_2$ .



white matter lesions on CT and MRI was gradually developed in small and repeated interobserver studies; the results with a more sensitive scale (four grades for different regions) have not been included in this report, because of considerable disagreement. We finally distinguished two regions that are most frequently affected: one region around the anterior horns of the lateral ventricles, and another consisting of the white matter around the posterior part of the cella media and the posterior part of the centrum semiovale. White matter lesions in these two regions of interest were separately evaluated on three subsequent CT or MRI slices: one through the choroid plexus of the posterior horns, one through the cella media, and one through the centrum semiovale (fig 1). These three levels can be considered representative of the total white matter, since they include the largest cross-sectional areas.

#### *CT scan assessment*

CT scans of 24 patients were obtained from the file of the Dutch TIA-trial.<sup>30</sup> Patients are eligible for this trial when they have suffered a transient ischaemic attack or a non-disabling stroke (Rankin-scale grade 3 or less).<sup>31</sup> CT scans were selected according to the presence and degree of white matter lucencies, with approximately equal proportions between normal white matter, slight involvement and severe involvement in one or more regions. To avoid overrepresentation of good quality scans, second generation scans of eight patients were included, together with third generation scans in sixteen patients. Hard-copy images had been made, at a window width of 75 Hounsfield in 11 patients, and of 100 units in 13 patients. A lacunar infarct, defined as a sharply marginated low density lesion without mass effect, and within a recognised arterial territory, was present in nine scans, and an infarct of the cortical type in three.

Leukoencephalopathy on CT was distinguished from infarction by its ill-defined borders. The severity of hypodensity, if present, was expressed in one of two degrees for each of the two regions. In grade 1 the abnormality was restricted to the region adjoining the ventricles, in grade 2 the increased hypodensity involved the entire region from lateral ventricle to the

cortex. Eleven observers, six neurologists and five radiologists, were instructed and subsequently asked to grade any abnormality of the white matter for each of the two regions in complete sets of all 24 CT scans. The degree of radiolucency was not explicitly graded. When asymmetric hypodensities were found, the most severely affected side was taken into account. Since the two regions almost make up the total white matter, the grades given to the regions can be added together, to give an overall value between 0 and 4.

#### *MRI assessment*

Twenty four MRIs, obtained from a study of elderly hypertensive patients, were selected to represent an even distribution according to the severity of white matter lesions, judged by increased signal-intensity on T2-weighted images. All MRI were performed with a 1.5 T Philips gyrosan. Multiple slice spin-echo (SE) sequences were performed with a repetition time (TR) of 2000 ms and an echo time (TE) of 50 and 100 ms, producing a T2-weighted image. A heavily T1-weighted image was produced with a TR of 1500 ms, an inversion time (TI) of 575 ms, and a TE of 30 ms. Five observers were asked to grade the severity of the white matter lesions in the same slices as for CT assessment, again separately for the anterior and posterior region, in one of three categories. Focal lesions were judged present when the diameter was 1 mm or more on the hard copy images (true dimensions approximately 2 mm). The severity of the lesions was graded according to their number and configuration: grade 0, no lesion or only a single one; grade 1, multiple focal lesions; grade 2, multiple confluent lesions scattered throughout the white matter. Since periventricular lesions nearly always accompanied focal or confluent lesions, we did not grade these separately. The grades for the two regions were considered separately as well as added together.

#### *Statistical analysis*

The degree of agreement among many observers was expressed by means of kappa statistics.<sup>32</sup> Kappa values were calculated for each CT or MRI, and for all possible combinations of the 11 and five observers, respectively.<sup>33</sup> For the grades with both regions added

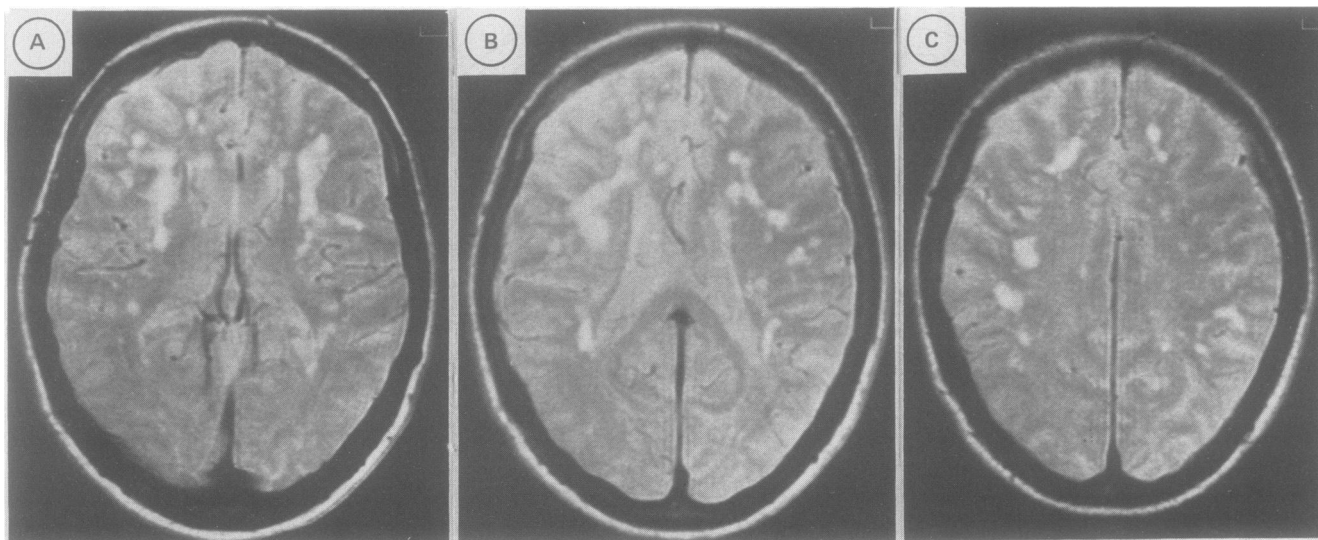


Figure 2 Multiple confluent lesions (grade 2) in the anterior white matter in the slice through the choroid plexus (left), and in the slice through the cella media (centre). Multiple focal lesions (grade 1) in the posterior white matter in the slice through the cella media (centre) and in the slice through the centrum semiovale (right).

together, kappa values were corrected by taking the degree of disagreement into account (weighted kappa).<sup>34</sup> In doing so, we attributed twice as much weight to a difference of two grades as to a difference of one grade, and three times as much weight to a difference of three grades (linear disagreement weights). Kappa values range between 0 (only chance agreement) and 1 (perfect agreement).

## Results

### CT assessment

With 11 observers, each set of scans of one patient gives rise to 55 pairwise comparisons, and with 24 scans the total number of comparisons is 1320. The kappa value for the anterior region alone was 0.51, and for the posterior region 0.57. When the grades of both regions were added together, the overall kappa value was 0.42, with a weighted kappa of 0.63 (table). One region was considered more affected than the other in 104 of 264 assessments.

On separate analysis of CT images made with the third generation scanners, the weighted kappa values were similar to those of the total series. In a further study with the same 24 CT scans, five of 11 observers were explicitly instructed to give separate grades for the anatomical extent of hypodensity and for the degree of radiolucency of the white matter; subsequently, they were asked to combine both features by attributing one grade. The interobserver agreement for the degree of radiolucency

was equal to that of the anatomical extent, and the combination of these separate features did not lead to better agreement than in the first round.

### MRI assessment

With five observers and 24 scans, the number of comparisons was 240. Perfect agreement for white matter lesions existed in 190 of these for the anterior region, and in 183 for the posterior region. The kappa value for the anterior region was 0.68, and for the posterior region 0.65. When the two grades were added together, the overall kappa value was 0.60, and the weighted kappa 0.78. One region was graded as more severely affected than the other in 34 of 120 assessments.

## Discussion

The reliability in measuring the severity of white matter lesions on CT and MRI has received little attention so far. An unequivocal grading scale is necessary for quantifying this abnormality, particularly in clinical and radiological studies involving large numbers of subjects, with cross-sectional or longitudinal design. The kappa values for assessments of the two regions with MRI (0.68 and 0.65) were higher than those for CT assessments of the same regions (0.51 and 0.57). The weighted agreements with the CT and MRI scale were 0.63 and 0.78, respectively, and these values can be considered good.<sup>35</sup>

That inter-observer agreement is lower for CT than for MRI may be mainly caused by the ill-defined margins of hypodensity on CT scan, which often makes the differences with normal white matter small. Another complicating factor with CT are partial volume effects with cerebrospinal fluid or grey matter.<sup>36</sup> Finally, lacunar infarcts often occurred in our series, as in other reports,<sup>37</sup> and these are sometimes difficult to distinguish from diffuse leukoencephalopathy, if the lacunes are only faintly seen on CT.

Table Kappa values for the severity of white matter lesions in anterior and posterior regions on CT and MRI

| CT                    | Kappa ( $K_w$ ) 95% confidence limits |                       |
|-----------------------|---------------------------------------|-----------------------|
| Anterior region       | 0.51                                  | 0.47–0.55             |
| Posterior region      | 0.57                                  | 0.53–0.60             |
| Both regions together | 0.42 (0.63)                           | 0.39–0.46 (0.61–0.66) |
| MRI                   |                                       |                       |
| Anterior region       | 0.68                                  | 0.60–0.76             |
| Posterior region      | 0.65                                  | 0.57–0.73             |
| Both regions together | 0.60 (0.78)                           | 0.52–0.67 (0.73–0.83) |

The severity of white matter lesions on MRI showed considerable variation. The greater image contrast of these lesions with the surrounding normal brain allows more reliable assessment. Moreover, exact measurement of the size of these lesions is possible. In two recent small series of patients,<sup>24,26</sup> the diameter of these lesions has been measured and added together. The three standardised axial levels for assessment in this study also offer the opportunity of measuring the cross-sectional areas of MRI lesions, and of detecting changes in longitudinal studies.

We found that the lesions in anterior or posterior white matter were not equally severe in about a third of the CTs and MRIs in our series. Also, the involvement of certain deep subregions, such as the internal or external capsule, is a conspicuous feature in some cases. In two CT studies, the extent of abnormal lucencies were assessed in several regions, resulting in a very sensitive summed score.<sup>7,12</sup> In one MRI study, the size of the lesions has been measured in four regions and added together.<sup>26</sup> However, this method results in greater interobserver disagreement, and it makes the scale impractical for use in large series of patients. It is unknown whether the distribution of lesions between anterior and posterior regions is only a matter of chance or whether it has pathogenetic significance. In view of the conflicting results about the existence of neuropsychological impairment in patients with white matter lesions, the distinction in anterior and posterior regions may be helpful.

In conclusion, the use of the three-point scale for white matter lesions on CT and MRI is suitable for analysing large series of patients. The three-point rating scale for MRI lesions in two regions has already been fruitfully applied in a recently completed study of elderly hypertensive patients.

We gratefully acknowledge the cooperation of neurologists and radiologists from the University Hospital, Utrecht, and of radiologists from the Academic Medical Centre, Amsterdam, from the Dijkzigt Hospital and St Franciscus Hospital, Rotterdam.

- 1 Valentine AR, Moseley IF, Kendall BE. White matter abnormality in cerebral atrophy: clinico-radiological correlations. *J Neurol Neurosurg Psychiatry* 1980;43:139-42.
- 2 Zeumer H, Schonsky B, Sturm KW. Predominant white matter involvement in subcortical arteriosclerotic encephalopathy (Binswanger disease). *J Comput Assist Tomogr* 1980;1:39-48.
- 3 Goto K, Ishii N, Fukasawa H. Diffuse white matter disease in geriatric population. *Radiology* 1981;141:687-95.
- 4 Brott Th, Mouradin M, Uhtman B. Dementia, ventricular enlargement, and hypertension associated with periventricular hypodensity by computed tomography. *Ann Neurol* 1986;20:129-30.
- 5 Loizou LA, Kendall BE, Marshall J. Subcortical arteriosclerotic encephalopathy: clinical and radiological investigation. *J Neurol Neurosurg Psychiatry* 1981;44:294-304.
- 6 Derix MMA, Hijdra A, Verbeeten BWJ. Mental changes in subcortical arteriosclerotic encephalopathy. *Clin Neurol Neurosurg* 1987;89:71-78.
- 7 Erkinjuntti T, Ketonen L, Sulkava R, Sipponen J, Vuorilaho M, Iivonen M. Do white matter changes on MRI and CT differentiate vascular dementia from Alzheimer's disease? *J Neurol Neurosurg Psychiatry* 1987;50:37-42.
- 8 Johnson KA, Davis KR, Buonanno FS, Brady ThJ, Rosen J,

- Growdon JH. Comparison of magnetic resonance and roentgen ray computed tomography in dementia. *Arch Neurol* 1987;44:1075-80.
- 9 George AE, de Leon MJ, Kalnin A, Rosner L, Goodgold A, Chase N. Leukoencephalopathy in normal and pathologic aging: 2. MRI of brain lucencies. *AJNR* 1986;7:567-70.
- 10 George AE, de Leon MJ, Gentes CI, et al. Leukoencephalopathy in normal and pathologic aging 1. CT of brain lucencies. *AJNR* 1986;7:561-6.
- 11 Steingart A, Hachinski VC, Lau C, et al. Cognitive and neurologic findings in demented patients with diffuse white matter lucencies on computed tomographic scan (Leuko-araiosis). *Arch Neurol* 1987;44:36-9.
- 12 Rezek DL, Morris JC, Fulling KH, Gado MH. Periventricular white matter lucencies in senile dementia of Alzheimer type and in normal aging. *Neurology* 1987;37:1365-8.
- 13 Inzitari D, Diaz F, Fox A, et al. Vascular risk factors and leuko-araiosis. *Arch Neurol* 1987;44:42-7.
- 14 Awad IA, Spetzler RF, Hodak JA, Awad CA, Carey R. Incidental subcortical lesions identified on magnetic resonance imaging in the elderly. 1. Correlation with age and cerebrovascular risk factors. *Stroke* 1986;17:1084-9.
- 15 Gerard G, Weisberg LA. MRI periventricular lesions in adults. *Neurology* 1986;36:998-1001.
- 16 Lechner H, Schmidt R, Bertha G, Justich E, Offenbacher H, Schneider G. Nuclear magnetic resonance image white matter lesions and risk factors for stroke in normal individuals. *Stroke* 1988;19:263-5.
- 17 Lotz PR, Ballinger WE, Quisling RG. Subcortical arteriosclerotic encephalopathy: CT spectrum and correlation. *AJNR* 1986;7:817-22.
- 18 Awad IA, Johnson PC, Spetzler RF. Incidental subcortical lesions on magnetic resonance imaging in the elderly. 2. Postmortem pathological correlations. *Stroke* 1986;17:1090-7.
- 19 Aharon-Peretz J, Cummings JL, Hill MA. Vascular dementia and dementia of the Alzheimer type. Cognition, ventricular size, and leuko-araiosis. *Arch Neurol* 1988;5:719-21.
- 20 Gupta SR, Naheedy MH, Young JC, Ghobrial M, Rubino FA, Hindo W. Periventricular white matter changes and dementia. Clinical, neuropsychological, radiological and pathological correlation. *Arch Neurol* 1988;45:637-41.
- 21 Kinkel WQR, Jacobs L, Polachini I, Bates V, Heffner RR. Subcortical arteriosclerotic encephalopathy (Binswanger's disease). *Arch Neurol* 1985;42:951-9.
- 22 Kertesz A, Black SE, Tokar G, Benke Th, Carr T, Nicholson L. Periventricular and subcortical hyperintensities on magnetic resonance imaging "Rims, caps and unidentified bright objects". *Arch Neurol* 1988;45:404-8.
- 23 Hershey LA, Modic MT, Greenough PG, Jaffe DF. Magnetic resonance imaging in vascular dementia. *Neurology* 1987;37:29-35.
- 24 Rao SM, Mittenberg W, Bernardin L, Haughton V, Leo GJ. Neuropsychological test findings in subjects with leuko-araiosis. *Arch Neurol* 1989;46:40-4.
- 25 Hendrie HC, Farlow MR, Guerriero Austrom M, Edwards MK, Williams MA. Foci of increased T2 signal intensity on brain scans of healthy elderly subjects. *AJNR* 1989;10:703-7.
- 26 Hunt AL, Orrison WW, Yeo RA, et al. Clinical significance of MRI white matter lesions in the elderly. *Neurology* 1989;39:1470-4.
- 27 Steingart A, Hachinski V, Lau C, Fox AJ, et al. Cognitive and neurologic findings in subjects with diffuse white matter lucencies on computed tomographic scan (leuko-araiosis). *Arch Neurol* 1987;44:32-5.
- 28 Lee A, Yu YL, Tsoi M, Woo E, Chang CM. Subcortical arteriosclerotic encephalopathy—a controlled psychometric study. *Clin Neurol Neurosurg* 1989;91:235-40.
- 29 Lee D, Fox A, Vinuela F, Pelz D, Lau C, Donald A, Meiskey H. Interobserver variation in computed tomography of the brain. *Arch Neurol* 1987;44:30-1.
- 30 The Dutch TIA study group. The Dutch TIA trial: the protective effect of low dose aspirin and atenolol in patients with transient ischemic attacks or non-disabling stroke. *Stroke* 1988;19:512-17.
- 31 van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJA, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988;19:604-7.
- 32 Longstreth WT, Koepsell TD, van Belle G. Clinical neuroepidemiology 1. diagnosis. *Arch Neurol* 1987;44:1091-499.
- 33 Schouten HJA. Measuring pairwise interobserver agreement when all subjects are judged by the same observers. *Statist Neerland* 1982;36:45-61.
- 34 Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1976;70:213-20.
- 35 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- 36 Schwartz M, Creasey H, Grady CL, DeLeo JM, Frederickson HA, Cutler NR, Rapoport SI. Computed tomographic analysis of brain morphometrics in 30 healthy men, aged 21 to 81 years. *Ann Neurol* 1985;17:146-57.
- 37 Bogousslavsky J, Regli F, Uske A. Leukoencephalopathy in patients with ischemic stroke. *Stroke* 1987;18:896-9.