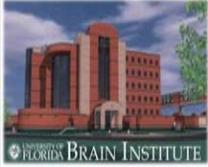


Quantification of White Matter Alterations: A Reliability Analysis



Catherine Price^{1,2}, Ilona Schmalfluss^{2,3,4}, Christopher Siström³
 University of Florida Departments of Clinical & Health Psychology¹, Anesthesiology², Radiology³,
 North Florida-South Georgia Veteran Administration Radiology Department⁴

RATIONALE

White Matter Alterations (WMAs) are also termed ‘subcortical hyperintensities’, ‘white matter lesions’, and ‘unidentified bright objects’. These names typify their bright and diffuse appearance in the white matter on T2 weighted MRI images. The clinical significance of WMAs have become of increasing interest as WMAs increase in frequency with advancing age and are associated with a variety of disorders.

Problem: Published studies vary in the 1) type of MRI sequence used to visualize WMAs and 2) type of measurement scale. Most frequently used are T2 weighted images and visual rating scales.

Purpose: To examine differences in the reliability of raters’ WMA measurements according to

- 1) type of MRI sequence (Fluid Attenuated Inversion Recovery (FLAIR) versus conventional T2-weighted) and
- 2) measurement technique (semi-automated pixel thresholding versus traditional visual rating scales).

Hypothesis: Stronger (higher) rater reliabilities will be obtained when WMAs are measured on FLAIR images and with a semi-automated computerized thresholding technique.

PARTICIPANT SAMPLE

Following IRB approval, 140 out of 3,382 MRI scans obtained in a 12 month time frame were selected for rating using a set inclusion and exclusion criteria.

Inclusion criteria: age 60 or older, presence of WMA, T2 and FLAIR images available, adequate quality

Exclusion criteria: prior brain surgery, infarctions, tumor

In order to obtain a sample of brains with a range of WMA severity, all 140 brains were graded as mild, moderate, severe by a neuroradiologist (IS).

After removing header information, 15 of the 140 scans (5 mild, 5 moderate, 5 severe) were randomly selected, separated by sequence (FLAIR, T2) and randomized for repeat scoring.

PROCEDURES

Rater Procedures

Two raters blind to severity rank independently conducted 90 ratings of randomly organized unidentified scans.

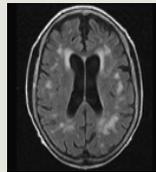
- 3 times for each FLAIR scan
- 3 times for the matching T2 scan

For these ratings, two technique were used:

- Pixel Thresholding Technique.
- Rotterdam Visual Rating Scale.

Type of MRI Sequences Rated

FLAIR

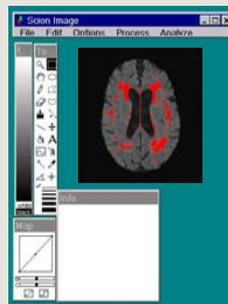


T2 weighted



Rating Measurement Techniques

Pixel Thresholding



Scion Image is a derivative of NIH Image. Both are image processing freewares (rsbweb.nih.gov).

Visual Rating Scale

	R	L
Periventricular Lesions (PV)		
Frontal Horns		
Lateral Wall		
Posterior Horns		
Total Score (0-9) = sum of most severe lesion in each location	calculation	calculation

PV lesions are characterized as:
 0 = no lesions
 1 = pencil thin
 2 = smooth halo
 3 = larger confluent lesion

	R	L
Subcortical and Deep WMA		
# of small lesions (1-3 mm)		
# of medium lesions (3-10mm)		
# of large subcortical lesions (>10 mm)		
total volume of lesions	calculation	calculation

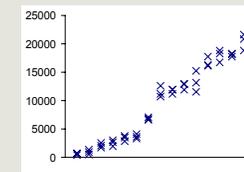
The total volume of the subcortical and / or deep white matter lesions is calculated with the assumptions that the lesions are spherical in appearance with an average diameter of 2, 6, and 12 mm for the 3 categories.

RESULTS

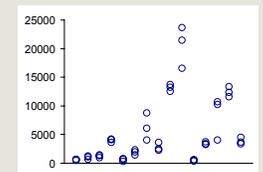
Intra-class correlation (ICC) coefficients

	mean	error variance	patient variance	reliability (ICC)
FLAIR_pixel	9487	0.455	48.5	0.99
FLAIR_visual	5366	2.49	33.7	0.93
T2_pixel	8989	1.76	41.0	0.96
T2_visual	7222	17.3	31.4	0.64

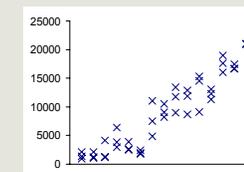
FLAIR – Pixel*



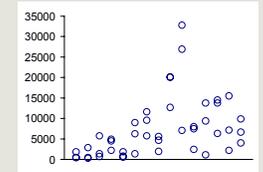
FLAIR – Visual*



T2 – Pixel*



T2 – Visual*



*Depicts individual rater scores for 3 time periods. The order in which the patients are shown is the SAME for each plot

CONCLUSION

- The Pixel Threshold method applied to FLAIR images is the most reliable approach for the analysis of WMAs.
- Moreover, the Pixel Threshold method provided high reliability within and between the T2 and FLAIR sequences. Consequently, both FLAIR and T2 images are equally appropriate to use when a reliable quantification method is applied.
- Applying a Visual Rating technique to T2 images yielded the lowest reliability.