Measuring longitudinal white matter changes: a comparison of a visual rating scale and a quantitative volumetric method

D. M. Van den Heuvel[1], A. C. van Es[1], V. H. ten Dam[2], W. M. Palm[1], F. Admiraal-Behloul[3], A. Spilt[1], G. Blauw[2], E. L. Bollen[4], R. G. Westendorp[2], M. A. van Buchem[1]

[1]radiology, Leiden University Medical Center, Leiden, Netherlands, [2]internal medicine, section of geriatrics and gerontology, Leiden University Medical Center, Leiden, Netherlands, [3]radiology, section of image processing (LKEB), Leiden University Medical Center, Leiden, Netherlands, [4]neurology, Leiden University Medical Center, Leiden, Netherlands

Introduction Up till now the progression of WML in longitudinal studies has mainly been estimated with use of visual rating scales.[1,2] However, the detection of longitudinal changes in WML load is problematic using these scales.[3] Particularly, the sensitivity of these scales can be disputed. Intra- and interrater reliability are low and the scales are hampered by ceiling and floor effects. Quantitative volumetric lesion detection methods offer a different method for obtaining WML load and might overcome these shortcomings. Compared to the visual WML ratings, quantitative WML measurements are more objective and reliable and provide us with exact measurements of WML volume. In this study we aimed to investigate the potential of both a widely used visual rating scale (i.e. the adapted Scheltens scale)[4] and a quantitative volumetric method to study longitudinal WM changes. WML load is associated with age[5]. Therefore, we reasoned that the most sensitive WML measurement in detecting longitudinal change would presumably translate well into a closer relationship with age. We hypothesize that the quantitative WML volume measurement shows stronger associations with age than the visual WML ratings.

Materials and methods The data in this study are drawn from the MRI substudy of the PROspective Study of Pravastatin in the Elderly at Risk (PROSPER).[6] MRI was performed on a clinical MR-system operating at a 1.5 Tesla field strength (Philips Medical Systems, Best, The Netherlands). Dual fast spin-echo images were obtained in all subjects. In 100 randomly chosen subjects WML load at baseline and follow up (mean 2.8 yrs) was evaluated with use of both the visual, semi-quantitative Schelten's scale (rank, 0 to 63) and an inhouse developed quantitative semi-automated WML detection software (WML volumes, cc). The agreement among observers for visual WML ratings and quantitative WML volume measurements was assessed by Intraclass Correlations Coefficients (ICC). Moreover, Spearman correlations of WML ratings and WML volumes with age were calculated. The level of statistical significance was set at $p<0.05$.

Results The mean age of the study group was 74.5 years (range, 70 to 81 years) and 41% were women. Reliability of visual WML ratings was good (ICC = .77 to .83) whereas reliability of quantitative WML measurements was excellent (ICC = .99 to 1). As expected, WML load at baseline was significantly associated with age. However, we found stronger associations for baseline quantitative WML measurements with age than for baseline visual WML ratings with age (r = .3, *p-value* = .002 and r = .2, *p-value* = .045, respectively). Moreover, a longitudinal evaluation of WML load showed that 27% of the subjects showed an unexpected regression in WML load when investigated with use of the visual rating scale whereas, in contrast, only 1% regressed when the quantitative method was used (figure 1a and 1b). Furthermore, the correlation of age with the quantitative WML progression estimate was twice as high compared to the progression estimate of the visual WML rating with age (r = .2, *p-value* = .035 and r = .4, *p-value* = .000 respectively).

Figure 1a: WML longitudinal evaluation: Visual rating scale

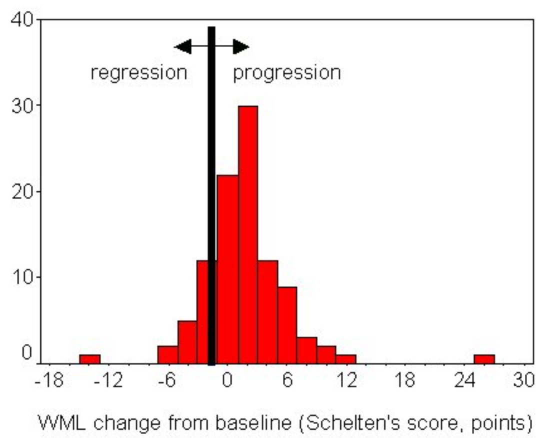
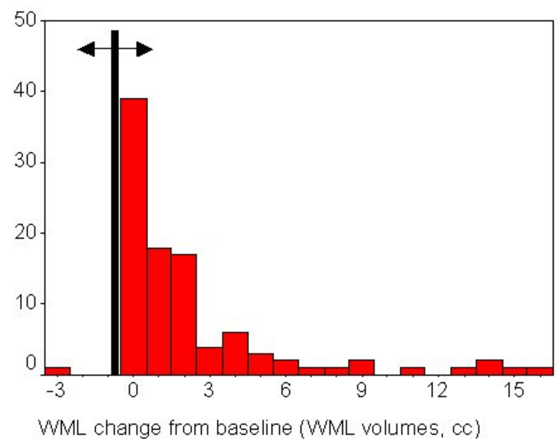


Figure 1b: WML longitudinal evaluation: Quantitative method



Conclusion Our data demonstrate that both the visual WML rating and quantitative WML volume measurement reflect the natural course of WML progression with age. However, the quantitative method, which provides us with more detailed information, seems to be more sensitive to depict these longitudinal WM changes. We suggest that quantitative volumetric assessment of WML load is more objective, reliable and exact and therefore better fit for assessment of longitudinal WM changes.

References

1. Schmidt R et al.Neurology 1999; 53(1):132-139.
2. Veldink JH et al.Neurology 1998; 51(1):319-320.
3. Kapeller P et al. Stroke 2003; 34(2):441-445.
4. Scheltens et al.: J. Neurol Sci 1993: 114:7-12.
5. Longstreth WT et al. Stroke 1996; 27(8):1274-1282.
6. Shepherd J et al. Am J Cardiol 1999; 84(10):1192-1197