Interobserver Agreement for Cerebral Glioma Volumetrics on Conventional MR Imaging

G. Thompson[1], J. R. Cain[1], A. Jackson[1], and S. J. Mills[1]

[1]Imaging Sciences and Biomedical Engineering, University of Manchester, Manchester, Greater Manchester, United Kingdom

Introduction: Cerebral gliomas are routinely investigated using conventional MRI. Subjective and qualitative assessment of these images are used to guide diagnosis, prognosis and treatment decisions, whilst treatment response and disease progression are often assessed on uni- or bi-dimensional manual quantitative measures such as the RECIST criteria [1]. Tumour volume characteristics have been shown to relate both to tumour subtype and prognosis, and thus have been suggested as a stratification tool for clinical trial design. Simple 1D and 2D measurements on T_2-weighted imaging relate poorly to tumour volume and prognosis in recurrent gliomas, the relationship with T_1-weighted measurements being less clear [1-4]. Many novel imaging biomarkers are under investigation to aid diagnosis, subtyping, outcome prediction and monitor follow-up. These can, however, require additional scanning and analysis time, and specialised software, equipment or skills not currently in widespread clinical use. When assessed quantitatively, the rich 3D data set obtained from routine conventional imaging may provide further diagnostic, prognostic and treatment response information. This will be of value, however, only if these measures are easy to perform with good interobserver agreement [5]. This study aims to investigate whether a straightforward method for assessment of both low and high grade gliomas has sufficient interobserver agreement to allow routine measurement of tumour volume, enhancing volume and T_2 bright volume (and, therefore, derived measures of enhancing fraction and peritumoural oedema volume.)

Methods: 29 patients with cerebral glioma (18 grade IV, 12 grade II; one having two synchronous grade II tumours of distinct histological subtype) were imaged on a 3.0T Philips Achieva scanner prior to surgery. Routine imaging included T_1WIR, T_2, FLAIR and post contrast T1 imaging. Independent retrospective analysis was conducted by two non-specialist radiologists (GT and JRC) blinded to the histological subtype. Using a pen tablet, regions of interest (ROI's) were drawn manually (Figure 1a) on each scan slice using MRIcro (http://www.sph.sc.edu/comd/rorden/mricro.html) and the sum of these ROI's used to determine the volume of interest (VOI). The high grade enclosed enhancement or low grade solid tumour volume (V_{T1C}) was defined post contrast. Software intensity filtering within MRIcro was used to determine the volume of enhancing tumour (V_E) within the V_{T1C} in a semi-automated manner. The T_2 bright volume was manually circumscribed on T_2W or FLAIR imaging (V_{T2}). Interobserver agreement was assessed for each measurement variable using the graphical method described by Bland and Altman [6], and by deriving Intraclass Correlation Coefficients (ICC; two-way random model with absolute agreement and 95%CI) using SPSS (SPSS Inc., Chicago, IL) [7].

Results: Average time for analysis was less than 10 minutes per tumour. Acceptable agreement between the two observers was seen both graphically and numerically for each volume parameter (ICC>0.94, Cronbach's α >0.96, F>2 for df=n-1, and p<0.001). Table 1 includes the values describing the level of interobserver agreement. Figure 1 shows examples of: b) bivariate comparison of observers showing line of equality; c) a histogram of interobserver differences with superimposed derived distribution curve demonstrating normality; and d) interobserver difference versus mean with 95% limits of agreement.

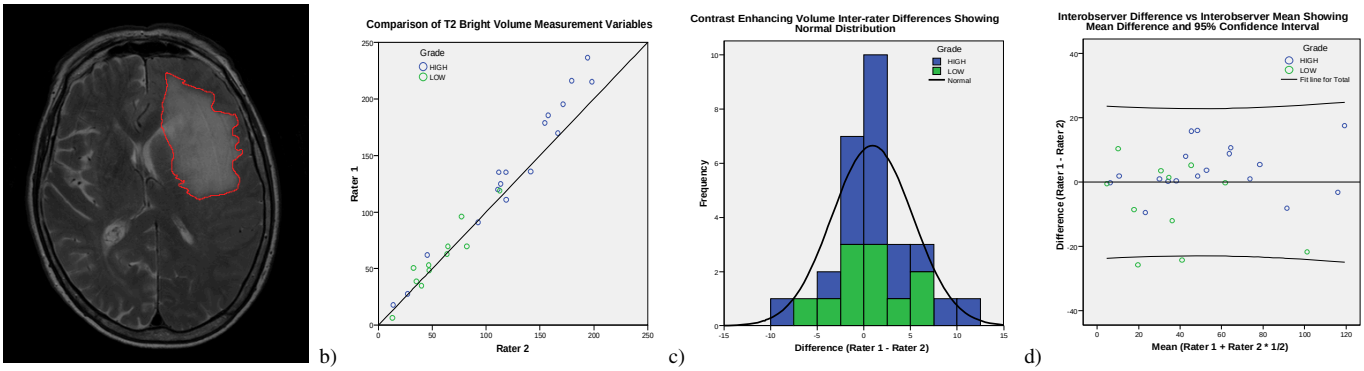

Figure 1: a) T_2W axial slice showing manual ROI. b) V_{T2} bivariate comparison showing line of equality (x=y). c) V_E interobserver difference frequency histogram showing normal distribution (verified with P-P and Q-Q plots, not shown.) d) V_{T1C} interobserver difference versus observers' mean with 95% Limits of Agreement. (High grade gliomas are shown in blue; low grade gliomas in green throughout.)

Variable	*n*	*Within LoA**	r^2	*ICC***	*ICC 95% CI*	α	*F-test (n-1)*	*p*
V_{T1C}	29	89.7%	0.882	**0.941**	0.878 to 0.972	0.968	31.7	<0.001
V_E	29	92.9%	0.917	**0.957**	0.912 to 0.980	0.978	46.2	<0.001
V_{T2}	28	96.4%	0.971	**0.964**	0.858 to 0.987	0.987	79.9	<0.001

Table 1: Interobserver agreement for manual measurement of T1 tumour volume (V_{T1C}) and T2 bright volume (V_{T2}); with semi-automated measurement of enhancing tumour volume (V_E). *Limits of Agreement (95% CI of mean differences.) **1=perfect correlation; 0=no correlation other than by chance.

Discussion: New imaging techniques are continually under investigation for their diagnostic and prognostic power in glioma. These often require extra time, skill, and equipment to process the resultant data. Current conventional imaging provides a rich 3D dataset which may yet provide additional diagnostic and prognostic information. The method described here for measurement of tumour volume, enhancing volume and T_2 bright volume shows acceptable interobserver agreement, and is both practical and practicable using freely available software, without the need for additional observer training.

References

1. Galanis, E., et al. Neuro Oncol, 2006. **8**(2): p. 156-65.
2. Dempsey, M.F., B.R. Condon, and D.M. Hadley. AJNR Am J Neuroradiol, 2005. **26**(4): p. 770-6.
3. Chow, K.L., et al. AJNR Am J Neuroradiol, 2000. **21**(3): p. 471-8.
4. Aghi, M., et al. Clin Cancer Res, 2005. **11**(24 Pt 1): p. 8600-5.
5. Joe, B.N., et al. Radiology, 1999. **212**(3): p. 811-6.
6. Bland, J.M. and D.G. Altman. Ultrasound Obstet Gynecol, 2003. **22**(1): p. 85-93.
7. Shrout, P.E. and J.L. Fleiss. Psychological Bulletin, 1979. **86**: p. 420-428.