SHORT REPORT

Automated brain volumetrics in multiple sclerosis: a step closer to clinical application

C Wang,[1,2] H N Beadnall,[2] S N Hatton,[1,2] G Bader,[3] D Tomic,[3] D G Silva,[3]
M H Barnett[1,2]

[1]Sydney Neuroimaging Analysis Centre, Sydney, New South Wales, Australia
[2]Brain and Mind Centre, The University of Sydney, Sydney, New South Wales, Australia
[3]Novartis Pharma AG, Basel, Switzerland

Correspondence to
Associate Professor,
MH Barnett,
Sydney Medical School, Brain and Mind Centre, The University of Sydney, Sydney, NSW 2006 Australia; michael@sydneyneurology.com.au

ABSTRACT

Background Whole brain volume (WBV) estimates in patients with multiple sclerosis (MS) correlate more robustly with clinical disability than traditional, lesion-based metrics. Numerous algorithms to measure WBV have been developed over the past two decades. We compare Structural Image Evaluation using Normalisation of Atrophy-Cross-sectional (SIENAX) to NeuroQuant and MSmetrix, for assessment of cross-sectional WBV in patients with MS.

Methods MRIs from 61 patients with relapsing-remitting MS and 2 patients with clinically isolated syndrome were analysed. WBV measurements were calculated using SIENAX, NeuroQuant and MSmetrix. Statistical agreement between the methods was evaluated using linear regression and Bland-Altman plots. Precision and accuracy of WBV measurement was calculated for (1) NeuroQuant versus SIENAX and (2) MSmetrix versus SIENAX.

Results Precision (Pearson's r) of WBV estimation for NeuroQuant and MSmetrix versus SIENAX was 0.983 and 0.992, respectively. Accuracy (Cb) was 0.871 and 0.994, respectively. NeuroQuant and MSmetrix showed a 5.5% and 1.0% volume difference compared with SIENAX, respectively, that was consistent across low and high values.

Conclusions In the analysed population, NeuroQuant and MSmetrix both quantified cross-sectional WBV with comparable statistical agreement to SIENAX, a well-validated cross-sectional tool that has been used extensively in MS clinical studies.

INTRODUCTION

Brain volume loss (BVL) in patients with relapsing-remitting multiple sclerosis (RRMS) has been shown to occur at a faster rate (0.5–1.35% per year) compared with the BVL in age-matched healthy individuals (0.1–0.3% per year).[1 2] The average patient with MS receiving first-generation disease-modifying treatment (DMT) or no DMT loses approximately 0.7% BV per year.[2] BVL correlates with, and is emerging as a key predictor of, worsening future disability and cognitive impairment in patients with MS.[1 3]

Numerous manual, semiautomatic and fully automated algorithms to measure whole brain volume (WBV) from MRIs have been developed over the past two decades.[4] Structural image evaluation, using normalisation, of atrophy—cross-sectional (SIENAX) is a freely available software tool (part of the FMRIB Software Library (FSL); http://www.fmrib.ox.ac.uk) that is widely used by expert MRI

reading centres to measure cross-sectional WBV in MS clinical studies.[4] SIENAX uses a segmentation-based algorithm to measure cross-sectional WBV from a single MRI, which is then normalised to a standard space to yield a normalised WBV.[4]

In routine clinical practice, disease monitoring is gradually evolving to include BVL, which has been validated as a predictor of treatment response in patients with MS,[1 5] in addition to clinical and traditional, lesion-based MRI outcomes.

SIENAX is widely used in cross-sectional MS studies,[4] but application in routine clinical practice is limited by the requirement for trained personnel and specialised reading centres to perform accurate measurements.[6]

Recent advances in computational methods have made it desirable to aim for algorithms that are fast, fully automated (minimal observer dependency), accurate, reproducible and applicable in both the clinical trial setting and routine clinical practice.[7]

NeuroQuant (CorTechs Labs) is a fully automated software, approved by the FDA and CE-marked (http://www.wholebrainatrophy.com) for measuring cross-sectional brain volumes. Originally designed for quantification of Alzheimer disease neuroimaging data sets,[8] NeuroQuant software takes a high-resolution three-dimensional (3D) T1-weighted sagittal, non-contrast-enhanced MRI and automatically produces segmentation-based measurements of cortical and subcortical volumes. Briefly, the algorithm corrects for gradient non-linearity and field inhomogeneity, deletes non-brain tissue using an active contour model, and segments anatomic structures by non-linearly registering the volumes to an inbuilt probabilistic atlas and applying an iteratively derived neuroanatomic label to each voxel. NeuroQuant compares the volumes to a normative database adjusted for age, gender and intracranial volume.[9] Using NeuroQuant MRI volumetrics, Neilsen et al, reported that brain atrophy correlates with physical disability in patients with MS treated with DMTs in routine clinical settings.[10]

MSmetrix (IcoMetrix) is a fully automated, CE-marked (http://www.icometrix.com) proprietary method that performs unsupervised tissue and lesion segmentation using 3D T1-weighted and fluid-attenuated inversion recovery (FLAIR) MRIs.[7] Briefly, the FLAIR image is rigidly co-registered to the T1-weighted image; and probabilistic anatomical priors for grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF), defined in

Montreal Neurological Institute (MNI) space, are transferred to the T1-weighted image space using an affine registration followed by a non-rigid registration. An iterative algorithm based on probabilistic tissue priors is then used to segment the T1-weighted image into GM, WM and CSF, while maintaining spatial consistency, until convergence. Further, an iterative process that generates a map based on deviation of each voxel intensity from the combined tissue classes (outlier belief map) is used to segment WM lesions, which are then filled in on the bias-corrected T1-weighted image with their neighbourhood WM intensities. Jain et al[7] showed that MSmetrix accurately performed automatic segmentation and volumetry of brain lesions on two distinct MS data sets from two clinical centres. MSmetrix reportedly also uses a separate, registration-based method that incorporates a Jacobian integration technique to facilitate longitudinal brain volumetric analyses.[10]

Various algorithms have been compared for accuracy, precision and reproducibility, and their performance on different MRI sequences has been assessed in other publications.[11–13] In this analysis, we compared SIENAX to two fully automated software methods, NeuroQuant and MSmetrix, for use in the assessment of cross-sectional WBV in the patients with MS.

METHODS
Patients
We analysed scans from 61 patients with RRMS and 2 patients with clinically isolated syndrome fulfilling McDonald 2010 criteria for MS. Patients able to have MRI were consecutively recruited from a single MS clinic in Sydney, Australia. All patients provided informed consent.

MRIs
MRIs of patients were recorded using the GE Discovery MR750 3.0 T scanner with the following specifications: sagittal 3D-T1 inversion-recovery fast spoiled-gradient echo, repetition time (TR)=7.2 ms, echo time (TE)=2.8 ms, inversion time (TI) =450 ms, resolution=0.9 mm isotropic, flip angle=12, acquisition matrix=256×256, field of view (FOV)=230 mm^2 and average=1. Where relevant, scans were acquired at least 31 days after corticosteroid administration.

WBV determination
WBV measurements were calculated from MRIs using three different software programs, SIENAX, as described by Smith et al,[14] NeuroQuant, as described in http://www.cortechslabs.com/neuroquant/ and MSmetrix as described by Jain et al.[7] Thalamic volume was measured using FIRST, a model-based tool described by Patenaude et al.[15]

SIENAX pipeline modifications
Non-brain tissue was deleted from images by an expert neuroimaging analyst (CW) using the Brain Extraction Tool from the FMRIB Software Library (FSL) and manual cleaning prior to SIENAX analysis. To prevent bias due to WM lesions in segmentation-based WBV measurement, lesion in-painting was performed prior to SIENAX analysis. Briefly, the T2 MS lesion mask was delineated prior to lesion in-painting by an expert neuroimaging analyst (CW) with a semiautomatic region of interest tool (Jim, Xinapse System).

Thereafter, the lesion area on structural imaging was replaced with neighbouring normal-appearing WM using the lesion filling tool from FSL. No modifications were made to the automated NeuroQuant or MSmetrix pipelines.

Statistical evaluation of precision and accuracy
Statistical agreement between the methods was evaluated using linear regression and Bland-Altman plots as described by Bland and Altman.[16] However, because SIENAX was considered as the standard, the original Bland and Altman methods were slightly modified and SIENAX was reported on x-axes instead of the means of SIENAX and the (other) method to be compared.[17]

Statistical precision (r), assessed by means of the Pearson's coefficient; and accuracy (Cb), the difference between the new method and the reference of WBV measurements, were calculated for NeuroQuant versus SIENAX and MSmetrix versus SIENAX. p Values <0.05 were considered statistically significant. Analyses were performed with Stata (StataCorp. 2009. Stata Statistical Software: Release 11. College Station, Texas, USA: StataCorp LP.)

RESULTS
Baseline disease characteristics and MS treatment history
The baseline characteristics are presented in table 1.

In the 12 months prior to study enrolment, 20/63 patients were not receiving any DMT. Of the remaining 43/63 patients, 19/63 patients were treated with interferon β-1a (Avonex, 17 and Rebif, 2), 7/63 with interferon β-1b (Betaferon), 7/63 with glatiramer acetate (Copaxone), 5/63 with fingolimod (Gilenya), 4/63 with natalizumab (Tysabri) and 1 patient was receiving an unspecified DMT.

Comparison of NeuroQuant and SIENAX
WBV was estimated in NeuroQuant by adding the automatically reported values for brainstem, forebrain parenchyma, cerebellar GM and cerebellar WM volumes, as no specific WBV metric was generated by the software at the time of analysis. The statistical precision of WBV estimation for NeuroQuant versus SIENAX was 0.983 (p<0.001) (figure 1A) and the statistical accuracy was 0.871 (figure 1A). NeuroQuant showed a 5.5% volume 'overestimation' compared with SIENAX (figure 1A)

Table 1	Baseline characteristics
Variable	**All patients N=63***
Sex, n (%)	
Female	52 (82.5)
Male	11 (17.5)
Type of MS, n (%)	
RRMS	61 (96.8)
CIS	2 (3.2)
Age, mean±SD, years	38±9.8
Age at disease diagnosis, mean±SD, years	32±8.3
Disease duration, mean±SD, years	8±7.9
EDSS, mean±SD	2±1.5
SIENAX absolute BV, mean±SD (minimum–maximum), mL	1101±109.0 (853–1393)
NeuroQuant absolute BV, mean±SD (minimum–maximum), mL	1162±118.0 (902–1468)
MSmetrix absolute BV, mean±SD (min–max), mL	1107±103.0 (881–1364)
T2 lesion number, mean±SD (minimum–maximum)	51±38.2 (4–159)
T2 lesion volume, mean±SD (min–max), mL	7±8.2 (0.1–32)

*Three participants were excluded from the MSmetrix analysis.
BV, brain volume; CIS, clinically isolated syndrome; EDSS, Expanded Disability Status Scale; MS, multiple sclerosis; RRMS, relapsing-remitting multiple sclerosis; SIENAX, Structural Image Evaluation using Normalisation of Atrophy-Cross-sectional.
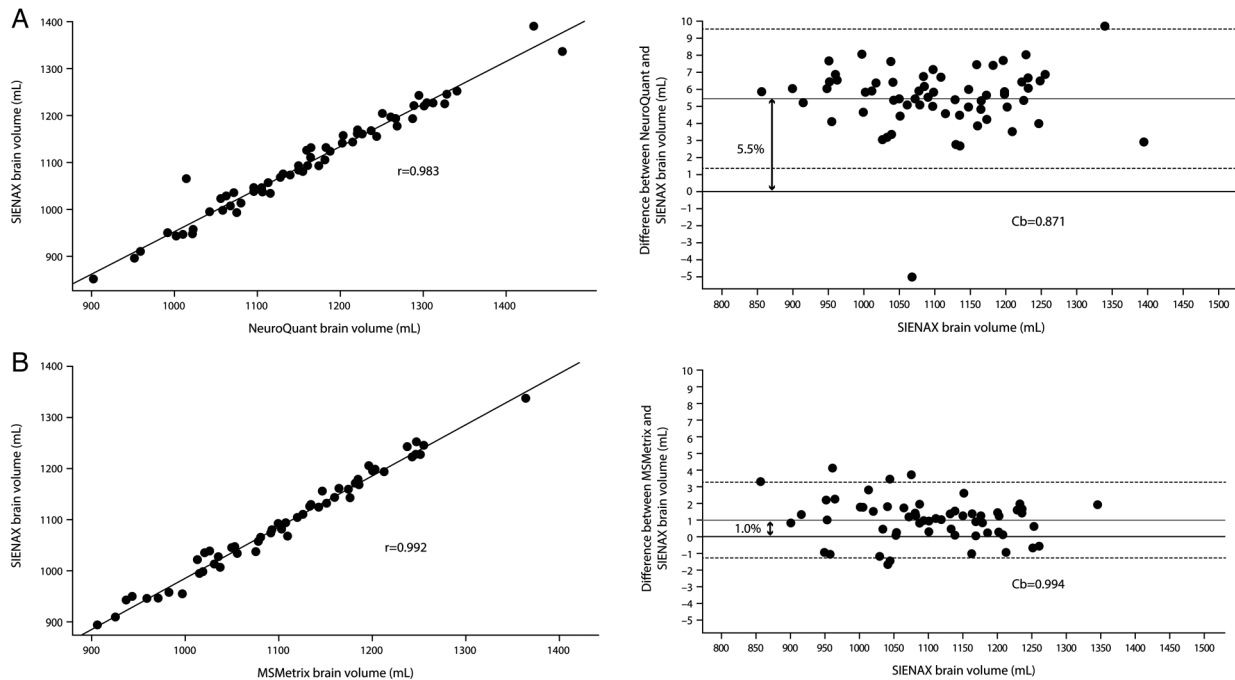
Figure 1 Whole brain volume measurement: comparison of precision and accuracy among three methods. (A) Comparison of NeuroQuant and structural image evaluation, using normalisation, of atrophy – cross-sectional (SIENAX). NeuroQuant showed a 5.5% volume 'overestimation' compared with SIENAX. (B) Comparison of MSmetrix and SIENAX. MSmetrix showed a 1% volume 'overestimation' compared with SIENAX.

that held across high and low volumes. Visual inspection of the NeuroQuant analysis for the single outlier shown in figure 1A revealed incomplete segmentation of a significant region of frontal GM. The statistical precision and accuracy of thalamic volume estimation for NeuroQuant versus FIRST were 0.786 (p<0.001) and 0.915, respectively. NeuroQuant showed a 6.3% volume 'overestimation' compared with FIRST for thalamic volume measurement. SIENAX and NeuroQuant-derived WBV both correlated inversely with T2 lesion volume (r=−0.43, p<0.001 and r=−0.41, p<0.001, respectively) and T2 lesion number (r=−0.38, p<0.005 and r=−0.35, p<0.01, respectively).

Comparison of MSmetrix and SIENAX
Three participants failed automated analysis and were excluded. Of the remaining participants, the statistical precision and accuracy of WBV estimation for MSmetrix versus SIENAX were 0.992 (p<0.001) and 0.994, respectively (figure 1B). MSmetrix showed a 1.0% volume 'overestimation' compared with SIENAX (figure 1B) that held across high and low volumes. The available iteration of MSmetrix software does not estimate thalamic volume. MSmetrix-derived WBV correlated inversely with T2 lesion volume (r=−0.37, p<0.005) and T2 lesion number (r=−0.32, p<0.05).

DISCUSSION
Patients with MS with a lower-than-expected WBV are at a higher risk of future disability worsening.[1] Additionally, patients with MS and a larger 'brain reserve' can better endure MS disease burden without noticeable cognitive decline.[18] Recent work has shown that a personalised WBV threshold value, based on normalised WBV, age, sex, Expanded Disability Status Scale (EDSS), disease duration and T2 lesion volume, can be used to predict risk of future disability worsening and cognitive performance on Paced Auditory Serial Addition Test (PASAT).[19]

The results of these studies also suggest that DMTs, which have a proven effect on slowing BVL in RRMS, may be beneficial for delaying disability worsening and improving cognition, especially in patients with low WBV threshold values.[20]

The principal technical requisites for incorporating measures of brain volume in routine clinical practice include speed, accuracy, reproducibility and 'user-friendliness' of the relevant method. NeuroQuant and MSmetrix have the potential to address some of these challenges,[7–9] but it is important that these tools are validated against SIENAX, the de facto gold standard for cross-sectional measurement of brain volume employed in modern MS clinical studies. Our findings show that NeuroQuant and MSmetrix both quantified cross-sectional WBV with comparable statistical agreement to SIENAX. We used MRI sequences that are employed in routine clinical practice in our centre; notably, the use of 3D-T1 sequence parameters that differ from those recommended by NeuroQuant (http://www.cortechslabs.com/wp-content/uploads/2015/06/NeuroQuant-Scanner-Parameters.pdf) did not appear to adversely affect segmentation results at the group level. MSmetrix does not recommend specific acquisition sequence parameters. Occasional scans failed, or provided inaccurate automated analysis using the methods tested in our study; improved algorithms in subsequent iterations of both software packages are expected to minimise the rate of automated analysis failure. While the development of these precise, easily deployable cross-sectional methods represents a first step toward the integration of brain volumetrics into routine clinical practice, emerging automated registration-based tools that permit rapid longitudinal assessment of WBV change may better predict longer term disease trajectory and guide therapeutic decision-making in individual patients with early MS.

Independent of the software used for quantification of WBV, several factors may interfere with the evaluation of WBV change, particularly outside of clinical trials. Image quality is

critical to the reliability of MRI quantification analysis, especially for potential clinical applications. BVL in patients with MS is also observed during the first 6 months to 1 year of DMT (principally due to resolution of oedema and inflammation) and stabilises from the second year of treatment. This complicates the efficacy interpretation of DMTs.[1] Also, technical factors, such as imaging variability between sites and across MRI machines, tissue contrast ratio, signal to noise ratio, artefacts and resolution; and biological factors, including hydration state[1] and diurnal variation in WBV,[21] can influence the accuracy of WBV measurements. Patient comorbidities, such as smoking, also affect WBV.[22]

Future advances in postprocessing imaging technology that 'correct' for biological, technical and treatment-related factors will facilitate the transition of this biomarker from clinical trials to routine clinical practice. The utility of the fully automated cross-sectional measures described here, referenced to appropriate healthy control databases, should be further explored as a prognostic tool in MS.

REFERENCES

1 De Stefano N, Airas L, Grigoriadis N, et al. Clinical relevance of brain volume measures in multiple sclerosis. CNS Drugs 2014;28:147–56.
2 Vollmer T, Signorovitch J, Huynh L, et al. The natural history of brain volume loss among patients with multiple sclerosis: a systematic literature review and meta-analysis. J Neurol Sci 2015;357:8–18.
3 Jacobsen C, Hagemeier J, Myhr K, et al. Brain atrophy and disability progression in multiple sclerosis patients: a 10-year follow-up study. J Neurol Neurosurg Psychiatry 2014;85:1109–15.
4 Giorgio A, Battaglini M, Smith SM, et al. Brain atrophy assessment in multiple sclerosis: importance and limitations. Neuroimaging Clin N Am 2008;18:675–86, xi.
5 Giovannoni G, Turner B, Gnanapavan S, et al. Is it time to target no evident disease activity (NEDA) in multiple sclerosis? Mult Scler Relat Disord 2015;4:329–33.
6 Miller DH, Barkhof F, Frank JA, et al. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. Brain 2002;125 (Pt 8):1676–95.
7 Jain S, Sima DM, Ribbens A, et al. Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. Neuroimage 2015;8:367–75.
8 Brewer JB, Magda S, Airriess C, et al. Fully-automated quantification of regional brain volumes for improved detection of focal atrophy in Alzheimer disease. AJNR Am J Neuroradiol 2009;30:578–80.
9 Nielsen A, Kita M, Baker J, et al. P530 Predictor status of disease modifying therapy time-exposure to brain atrophy using NeuroQuant at ACTRIMS-ECTRIMS MSBoston 2014: Poster Sessions 2. Mult Scler 2014;20(1 Suppl):285–496.
10 Ribbens A, Cambron M, De Keyser J, et al. C-2223 Quantifying brain atrophy in clinical practice for MS patients: a feasibility study on the measurement error. Presented at ECR; 2015, Vienna, Austria.
11 Sormani MP, Rovaris M, Valsasina P, et al. Measurement error of two different techniques for brain atrophy assessment in multiple sclerosis. Neurology 2004;62:1432–4.
12 Zivadinov R, Grop A, Sharma J, et al. Reproducibility and accuracy of quantitative magnetic resonance imaging techniques of whole-brain atrophy measurement in multiple sclerosis. J Neuroimaging 2005;15:27–36.
13 Smith SM, Rao A, De Stefano N, et al. Longitudinal and cross-sectional analysis of atrophy in Alzheimer's disease: cross-validation of BSI, SIENA and SIENAX. Neuroimage 2007;36:1200–6.
14 Smith SM, Zhang Y, Jenkinson M, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage 2002;17:479–89.
15 Patenaude B, Smith SM, Kennedy DN, et al. A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage 2011;56:907–22.
16 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327:307–10.
17 Krouwer JS. Why Bland–Altman plots should use X, not $(Y+X)/2$ when X is a reference method. Stat Med 2008;27:778–80.
18 Sumowski JF, Rocca MA, Leavitt VM, et al. Brain reserve and cognitive reserve protect against cognitive decline over 4.5 years in MS. Neurology 2014;82:1776–83.
19 Sormani MP KL, Radue EW, Cohen J, et al. Defining brain volume cut-offs to predict disability progression in MS:an analysis of a large cohort of relapsing-remitting MS patients. PS12.3 presented at 2014 Joint ACTRIMS-ECTRIMS Meeting; September 2014. Mult Scler 2014;20((1 Suppl)):suppl 14–66.
20 Sormani MP, Kappos L, Cohen J, et al. Normalized brain volume predicts cognitive performance in MS: an analysis of a large cohort from fingolimod phase III studies (P7.284). Neurology 2015;84(14 Supplement).
21 Nakamura K, Brown RA, Narayanan S, et al, Alzheimer's Disease Neuroimaging Initiative. Diurnal fluctuations in brain volume: statistical analyses of MRI from large populations. Neuroimage 2015;118:126–32.
22 Kappus N, Weinstock-Guttman B, Hagemeier J, et al. Cardiovascular risk factors are associated with increased lesion burden and brain atrophy in multiple sclerosis. J Neurol Neurosurg Psychiatry 2016;87:181–7.