iMRI

Investigative Magnetic Resonance Imaging

Comparison of Vendor-Provided Volumetry Software and NeuroQuant Using 3D T1-Weighted Images in Subjects with Cognitive Impairment: How Large is the Inter-Method Discrepancy?

Jieun Chung[1], Hayoung Kim[1], Yeonsil Moon[2], Won-Jin Moon[1]

[1]Department of Radiology, Konkuk University School of Medicine, Seoul, Korea
[2]Department of Neurology, Konkuk University School of Medicine, Seoul, Korea

Original Article

Correspondence to:
Won-Jin Moon, M.D., Ph.D.
Department of Radiology, Konkuk University School of Medicine, 120-1 Neungdong-ro, Gwangjin-gu, Seoul 05030, Korea.
Tel. +82-2-2030-5499/5544
Fax. +82-2-2030-5549
E-mail: mdmoonwj@kuh.ac.kr or mdmoonwj@naver.com

Background: Determination of inter-method differences between clinically available volumetry methods are essential for the clinical application of brain volumetry in a wider context.

Purpose: The purpose of this study was to examine the inter-method reliability and differences between the Siemens morphometry (SM) software and the NeuroQuant (NQ) software.

Materials and Methods: MR images of 86 subjects with subjective or objective cognitive impairment were included in this retrospective study. For this study, 3D T1 volume images were obtained in all subjects using a 3T MR scanner (Skyra 3T, Siemens). Volumetric analysis of the 3D T1 volume images was performed using SM and NQ. To analyze the inter-method difference, correlation, and reliability, we used the paired t-test, Bland-Altman plot, Pearson's correlation coefficient, intraclass correlation coefficient (ICC), and effect size (ES) using the MedCalc and SPSS software.

Results: SM and NQ showed excellent reliability for cortical gray matter, cerebral white matter, and cerebrospinal fluid; and good reliability for intracranial volume, whole brain volume, both thalami, and both hippocampi. In contrast, poor reliability was observed for both basal ganglia including the caudate nucleus, putamen, and pallidum. Paired comparison revealed that while the mean volume of the right hippocampus was not different between the two software, the mean difference in the left hippocampus volume between the two methods was 0.17 ml (P < 0.001). The other brain regions showed significant differences in terms of measured volumes between the two software.

Conclusion: SM and NQ provided good-to-excellent reliability in evaluating most brain structures, except for the basal ganglia in patients with cognitive impairment. Researchers and clinicians should be aware of the potential differences in the measured volumes when using these two different software interchangeably.

Keywords: Brain volumetry; Reliability; Siemens morphometry, NeuroQuant

INTRODUCTION

Structural atrophy of specific brain regions is a valuable imaging marker for specific neurodegenerative dementia (1). Besides amyloid deposition, hippocampal atrophy can independently predict memory decline in non-demented subjects, and intracranial volume (ICV), and/or temporal lobe volume (2) can serve as the brain reserve that is beneficial in cognitive function (2, 3). Accordingly, volumetry of the hippocampus (HIP) and other brain regions has been incorporated into the clinical workup for memory and dementia (4). Also, quantitative volumetry is a valuable tool in monitoring otherwise healthy individuals wishing to evaluate their brain reserve in light of the epidemic era of dementia.

Currently, several commercially available clinical volumetry software are being studied (5-8). Diagnostic accuracy of these software has been extensively studied (4, 9-11). In patients with Alzheimer's disease (AD) and mild cognitive impairment (MCI), the diagnostic accuracy of hippocampal volumetry ranges 83-88% (4, 9-11). However, analytical accuracy parameters such as reliability, reproducibility, and measured bias have been evaluated in only a few studies (9, 12). Also, despite the surge of different clinical volumetry software from different developers, the lack of knowledge of their analytical accuracy raises concerns regarding their misuse or overuse by incognizant healthcare personnel.

NeuroQuant (NQ) is the first FDA-approved and the most commonly used clinical volumetry software, which is a spin-off of FreeSurfer, a research-oriented software for the volumetry purpose (5). Siemens morphometry (SM) software is one of the most recently introduced software that has been incorporated into the MRI system, instead of the separate use of the software on an independent workstation. It uses a statistical inference approach based on the Markov random field image models to reflect unbiased prior anatomical knowledge as well as image characteristics such as RF inhomogeneity and partial volume effects (13). Until recently, there has been no report on the inter-method difference between SM and NQ.

Hence, in this study, we evaluated the inter-method reliability and potential differences between SM and NQ in patients with cognitive impairment.

MATERIALS AND METHODS

This retrospective study received Institutional Review Board approval, and the requirement for written informed consent was waived because of the retrospective study design.

Subjects

Eighty-six consecutive patients with subjective or mild cognitive impairment (31 males and 55 females; age 52-88; mean age 72.90) who visited a memory clinic and underwent brain 3T MRI January-August 2018 were included in this study. Clinical diagnosis was made by a neurologist with 13 years' experience: 14 patients with subjective cognitive impairment (one male and 13 females; age 60-83; mean age 71), 33 with MCI (10 men and 23 women; age 52-84; mean age 71), 17 with AD (nine males and eight females; age 68-85; mean age 77), 10 with vascular dementia (VaD) (six males and four females; age 62-85; mean age 74), four with dementia with Lewy bodies (DLB) (four females; age 67-88; mean age 75), two with frontotemporal lobar degeneration (FTLD) (one male and one female; age 63-83; mean age 73), 1 with Parkinson's disease dementia (PDD) (one male; age 68), and five with insufficient neuropsychiatric evaluation or stroke (three males and two females; age 56-79; mean age 72). All diagnoses were based on clinical history, physical examination, and neuropsychiatric evaluation.

The diagnoses of MCI, dementia, AD, VaD, DLB, FTLD, and PDD were based on the criteria suggested by Petersen et al. (14), the Diagnostic and Statistical Manual of Mental Disorders (4th ed.) (15), the criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (15), the criteria suggested by the National Institute of Neurological Disorders and Stroke of the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (16), The Lewy Body Composite Risk Score (17), International Behavioural Variant FTD Criteria Consortium (18), and the 2007 Movement Disorder Society guidelines (19), respectively.

Image Acquisition

All subjects underwent MRI with a 3-T unit (Skyra 3T, Siemens, Germany) using a 20-channel head coil. The routine MRI protocol included the following sequences: axial and sagittal T1-weighted inversion recovery imaging (TR/TE, 2300/2.98; inversion time, 900 ms; section thickness,

iMRI

Inter-Method Difference of Volumetric Softwares | Jieun Chung, et al.

1 mm; matrix, 256 × 256); axial FLAIR imaging (TR/TE, 5000/393; inversion time, 1800 ms; section thickness, 1 mm; matrix, 256 × 256); axial susceptibility-weighted imaging (TR/TE, 29/20; section thickness, 2 mm; matrix, 512 × 256; flip angle, 15°); and sagittal T1-weighted volumetric Magnetization Prepared RApid Gradient Echo (MPRAGE) (TR/TE, 2300/2.98; inversion time, 900 ms; section thickness, 1 mm; matrix, 256 × 256; flip angle, 9°; FOV, 250 × 250 mm).
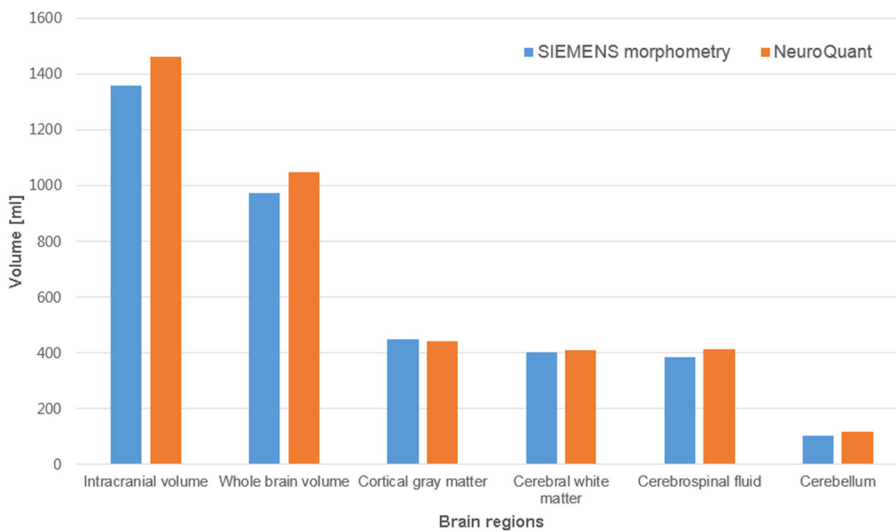
MR Volumetry

Sagittal T1-weighted volumetric images of patients with subjective and objective impairments were uploaded to the SM and NQ server, which provides computer-automated analysis of the brain images.

The processing in NQ was as follow: removal of the scalp, skull, and meninges; inflation of the brain to a spherical shape; mapping of the spherical brain to a common spherical space shared with the Talairach atlas coordinates; identification of the segmented brain regions; and deflation of the brain to its original shape.

SM processing involved the following steps: skull stripping; tissue classification to extract brain tissue compartments such as the white matter, gray matter, and intra/extra ventricular cerebrospinal fluid (CSF); checking of segmentation quality by correlation of the extracted gray matter map and asymmetry of the white matter map; segmentation of the central nuclei, HIP, brainstem, and ventricles; lobar parcellation; and detection of white matter abnormality.

Although SM and NQ provided the normative percentile compared to age- and sex-matched reference distribution, we only used the segmented volume of the specific brain regions for this study since the references of both software were not identical.

a
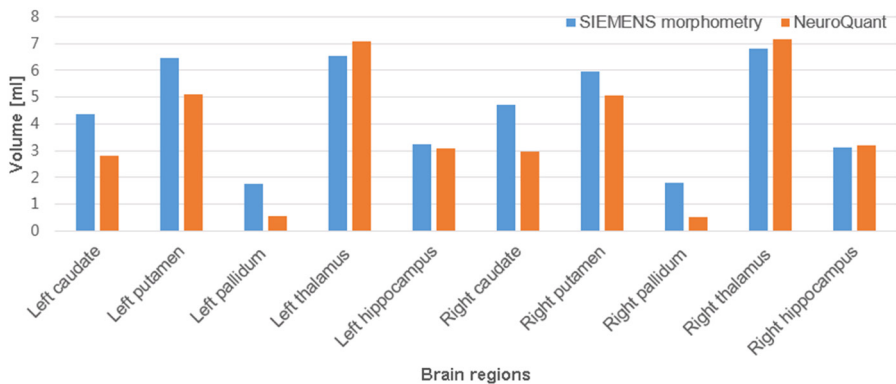
b

Fig. 1. Comparisons of volume measurements obtained using the Siemens morphometry and the NeuroQuant.

iMRI

Table 1. Comparison of Volume Measurements Obtained from NeuroQuant and Siemens Morphometry

	Siemens morphometry	NeuroQuant				
	Mean ± SD (ml)	Mean ± SD (ml)	P value			
Intracranial volume	1360.11 ± 124.75	1461.22 ± 127.58	<0.0001			
Whole brain volume	974.61 ± 93.16	1047.76 ± 94.03	<0.0001			
Cortical gray matter	449.28 ± 46.04	443.14 ± 42.9	0.0208			
Cerebral white matter	401.17 ± 49.27	410.9 ± 52.47	0.0229			
Cerebrospinal fluid	385.5 ± 75.59	413.47 ± 78.6	<0.0001			
Cerebellum	104.32 ± 11.59	118.37 ± 12.64	<0.0001			
	Left hemisphere			Right hemisphere		
	Siemens morphometry	NeuroQuant		Siemens morphometry	NeuroQuant	
	Mean ± SD (ml)	Mean ± SD (ml)	P value	Mean ± SD (ml)	Mean ± SD (ml)	P value
Caudate	4.38 ± 0.59	2.82 ± 0.64	<0.0001	4.73 ± 0.74	2.97 ± 0.73	<0.0001
Putamen	6.45 ± 0.72	5.12 ± 0.8	<0.0001	5.96 ± 0.66	5.07 ± 0.74	<0.0001
Pallidum	1.77 ± 0.24	0.55 ± 0.17	<0.0001	1.8 ± 0.21	0.52 ± 0.16	<0.0001
Thalamus	6.53 ± 0.7	7.1 ± 0.95	<0.0001	6.81 ± 0.78	7.16 ± 1.08	<0.0001
Hippocampus	3.24 ± 0.53	3.08 ± 0.62	0.0005	3.13 ± 0.51	3.18 ± 0.7	0.2959[†]

SD = standard deviation
[†]Statistically not significant

Statistical Analysis

To compare the agreement between volumes obtained by the SM and NQ, the paired t-test and Bland-Altman plot were used. The Pearson's correlation coefficient was calculated to measure the correlation between the two methods. Inter-method reliability between the SM and NQ was analyzed by two-way absolute intraclass correlation coefficient (ICC) with 95% confidence interval (CI). To interpret the ICC values, the following guidelines were used: poor reliability, ICC < 0.5; moderate reliability, 0.5 ≤ ICC < 0.75; good reliability, 0.75 ≤ ICC < 0.9; and excellent reliability, ICC ≥ 0.9 (20). The standardized mean difference between paired results of two software was evaluated through the effect size (ES). ES was defined as follow: trivial, ES < 0.2; small, 0.2 ≤ ES < 0.5; moderate, 0.5 ≤ ES < 0.8; and large, ES ≥ 0.8 (21).

P values < 0.05 indicated statistical significance. All statistical analyses were performed with statistical software packages (MedCalc version 18.2.1, MedCalc Software, Ostend, Belgium; SPSS, version 18 for Windows, SPSS, Chicago, IL, USA).

RESULTS

Paired t-test results for comparisons of volumes obtained using the SM and NQ are shown in Table 1 and Figure 1.

There were significant differences in the mean volumes of most brain regions, except for the right HIP (P = 0.296), between the two software. Compared with the NQ, SM showed significantly lesser ICV, whole brain volume (WBV), cerebral white matter (CWM), and CSF. In contrast, the SM showed larger volumes for most gray matter regions including cortical gray matter (CGM), caudate (CAU), putamen (PUT) and globus pallidus (GP), but not for the thalamus (THAL). Regarding the HIP, while right hippocampal volume was not different between the two methods, the volume of left HIP measured by SM was significantly larger than that measured by NQ. The Bland-Altman plot (Fig. 2) showed that the larger volume of HIP resulted in underestimation of the volume determined by the SM compared to that determined by NQ.

The Pearson's correlation coefficient between the SM and NQ showed significantly moderate to markedly strong correlation (0.6671 ≤ r ≤ 0.9640) in all measured structures. Regarding the inter-method reliability, excellent reliability was observed for CGM, CWM, and CSF. Good reliability was observed for ICV and WBV as well as some small structures

iMRI

Inter–Method Difference of Volumetric Softwares | Jieun Chung, et al.

such as THAL and HIP. Notably, ICCs of brain structures including both HIPs showed good correlation (ICC of left HIP, 0.8441, 95% CI: 0.7356-0.9046 and ICC of right HIP, 0.8417, 95% CI: 0.7575-0.8967). However, the inter-method reliability was poor for regions of the basal ganglia (both CAU, both GP, and the left PUT) (0.0334 ≤ ICC ≤ 0.3334).

To standardize the mean differences in the volumes of measured structures, we compared the ESs between the SM and the NQ. HIP measures showed trivial (right) to small ES (left). However, an undeniably large ES was observed in the basal ganglia regions and the cerebellum (Table 2).
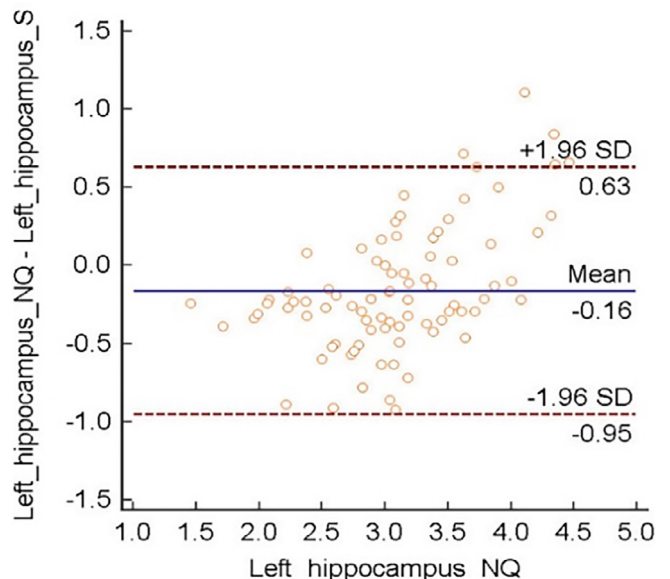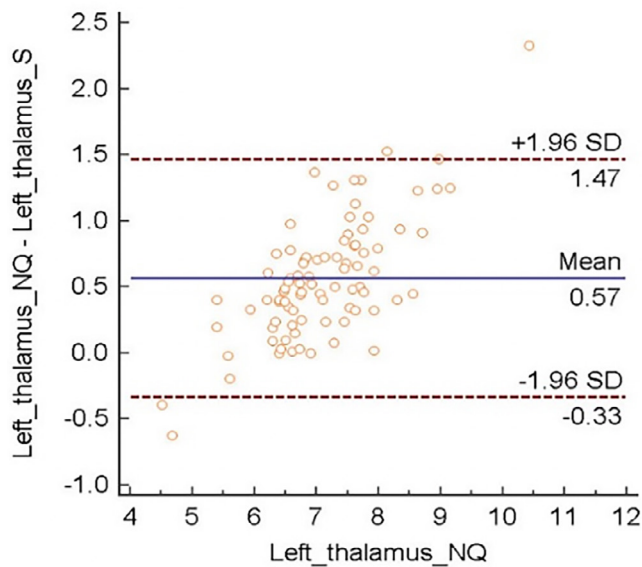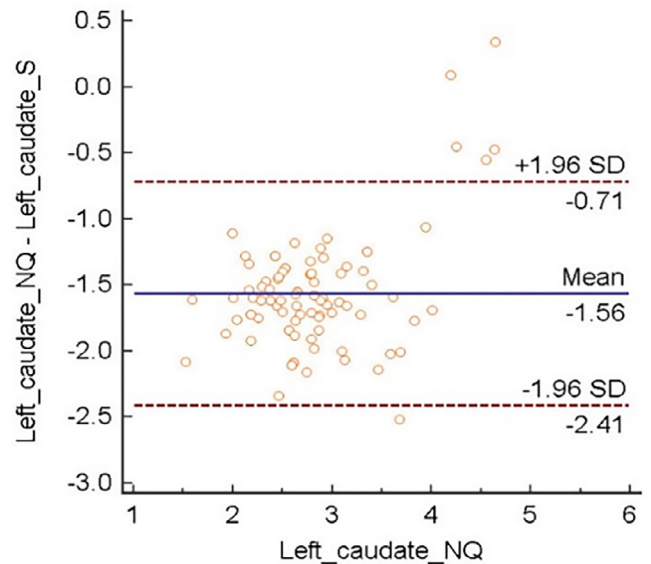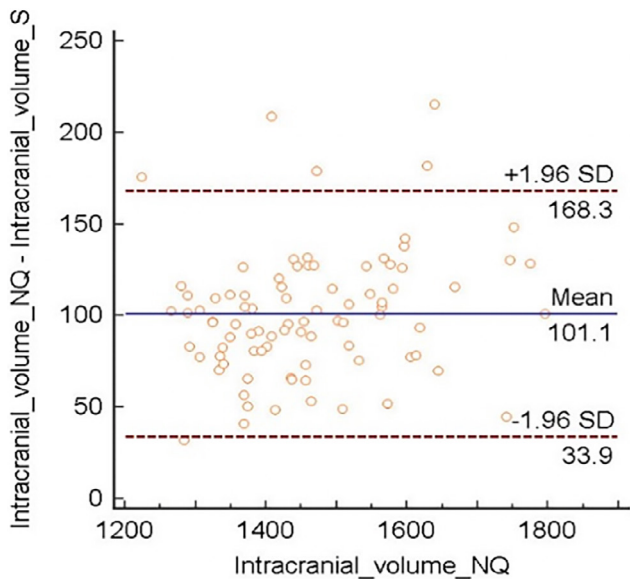
a

b

c

d

Fig. 2. The Bland-Altman plot showing the absolute difference in the intracranial volume, left caudate, left thalamus, and left hippocampus volumes measured by the Siemens morphometry and NeuroQuant against the absolute volume measured by the NeuroQuant. For a large volume structure such as the intracranial volume, the volume measured by the SM was smaller than that measured by the NQ (a). In contrast, for deep gray matter structures, the volume measured by the SM was larger than that by the NQ (b and d). For exception, the thalamus measured smaller by the SM as compared to the NQ (c).

Table 2. Results of Pearson's Correlation, Intraclass Coefficient, and Effect Size in Each Hemisphere

	r	P value	ICC (95% CI)	P value	Effect size (d')
Intracranial volume	0.9633	<0.0001	0.8434 (−0.0865 ~ 0.9622)	<0.001	0.7987
Whole brain volume	0.9640	<0.0001	0.8497 (−0.0873 ~ 0.9638)	<0.001	0.7813
Cortical gray matter	0.8543	<0.0001	0.9163 (0.8692 ~ 0.9461)	<0.001	−0.1371
Cerebral white matter	0.9396	<0.0001	0.9479 (0.7955 ~ 0.9783)	<0.001	0.2715
Cerebrospinal fluid	0.9389	<0.0001	0.9367 (0.5640 ~ 0.9781)	<0.001	0.3605
Cerebellum	0.9233	<0.0001	0.7124 (−0.0993 ~ 0.9211)	<0.001	1.1486

	Left hemisphere					Right hemisphere				
	r	P value	ICC (95% CI)	P value	Effect size (d')	r	P value	ICC (95% CI)	P value	Effect size (d')
Caudate	0.7394	<0.0001	0.2983 (−0.0749 ~ 0.6705)	<0.001	−2.5235	0.7760	<0.0001	0.3334 (−0.0718− 0.7065)	<0.001	−2.3999
Putamen	0.8351	<0.0001	0.4952 (−0.0994 ~ 0.8220)	<0.001	−1.7271	0.8124	<0.0001	0.6191 (−0.1813−0.8743)	<0.001	−1.2541
Pallidum	0.6671	<0.0001	0.0665 (−0.0199 ~ 0.2623)	<0.001	−5.5678	0.4297	<0.0001	0.0334 (−0.0203 ~ 0.1436)	<0.001	−6.7926
Thalamus	0.8899	<0.0001	0.8178 (0.0058 ~ 0.9385)	<0.001	0.5826	0.8393	<0.0001	0.8553 (0.6732 ~ 0.9244)	<0.001	0.3345
Hippocampus	0.7652	<0.0001	0.8441 (0.7356 ~ 0.9046)	<0.001	−0.2696	0.7609	<0.0001	0.8417 (0.7575 ~ 0.8967)	<0.001	0.0784

CI = confidence interval; ICC = intraclass correlation coefficient

DISCUSSION

This study compared two brain volumetry software, the SM and the NQ, in terms of their reliability in subjects with cognitive impairment. We found mostly good-to-excellent inter-method reliability and correlation for all brain structures, except for the basal ganglia. However, despite the high inter-method reliability, there was a significant difference in the measured volume of most regions, except for the right HIP, when using the paired-t-test. Additionally, basal ganglia and cerebellum showed an undeniably large ES that could lead to spurious results.

We found that larger structures showed smaller differences and higher reliability when measuring the brain volume by the SM and NQ. Compared to the ICV and WBV, CGM showed better ICC (excellent inter-method reliability). Our finding has a potential implication in estimating the brain reserve, which is the ability of the brain to tolerate aging and the pathology of dementia (2). ICV was initially suggested to be the brain reserve in non-dementia individuals (3). Recently, other researchers have suggested CGM as a potential marker for brain reserve (22).

In terms of the deep gray matter, there was a substantial difference in the volumes measured for the basal ganglia between the SM and NQ. The degree of poor inter-method reliability was severe for GP, CAU, and PUT, in that order. Our finding corroborated the previous inter-method comparison reports that basal ganglia volume measurements differ significantly between different tools (9, 23, 24).

Volumetric differences for HIP and THAL were also noted, but were rather small, thereby presenting good inter-method reliability. HIP volume can be used as an imaging marker for neurodegeneration related to the AD neuropathology and has been incorporated into the diagnostic framework (25). Thalamic volume measurement has been used as an adjunct marker for neurodegeneration in MS and other diseases (26). In a previous study, Schmitter et al. (27) reported different volumetric estimates for HIP between the FreeSurfer and MorphoBox in patients with MCI and AD. Our results support the use of volumetry of the HIP and THAL regardless of the platform or software used.

Regarding the apparent differences in the volumes of the basal ganglia, we presumed that the different atlas and segmentation models of the SM and NQ are the main reasons for these discrepancies (Fig. 3). Brain volume measurement by the NQ was comparable to that by the
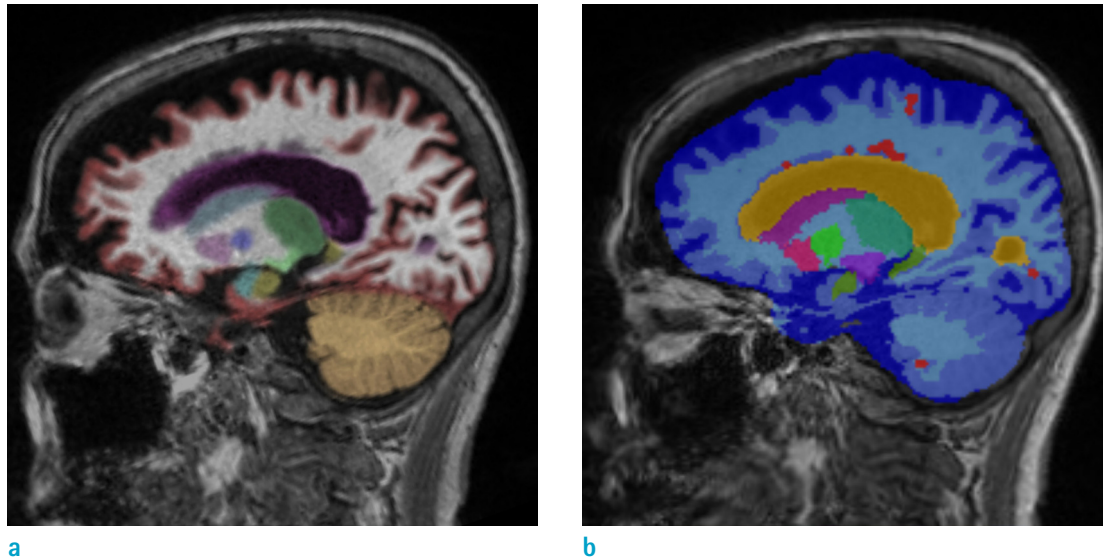
iMRI

Inter-Method Difference of Volumetric Softwares | Jieun Chung, et al.

a b

Fig. 3. The representative segmentation image for the Siemens morphometry and the NeuroQuant. (a) the Siemens morphometry; (b) the NeuroQuant.

FreeSurfer, a reference standard of volumetry (12, 23). The NQ uses a segmentation algorithm structure-wise similar to the FreeSurfer, but uses a different probabilistic atlas, an independent code base, methods for intensity normalization, and gradient distortion correction to accommodate scanner-specific acquisition-level differences. The SM is developed from the MorphoBox algorithm (28, 29). The MorphoBox prototype needs single-subject template instead of a prior atlas from several subjects for brain volumetry and applies a tissue-wise segmentation model (27).

Generally, volume estimates by the SM were smaller than those by the NQ. This systematic error is probably because of the different atlas and segmentation models of the two software and can be corrected by changing the MR parameters including spatial resolution, contrast, and filtering (30) and by using the reference values. Currently, the two software appear to apply different sets of normative database (28, 31). In future studies, a common normative database should be established.

Our study had limitations. First, our reference, the NQ, was not ground truth. True inter-method reliability can only be measured by a phantom study. Second, we did not use normative percentiles of volume measurements provided by each software. We believed that the use of normative values could potentially mitigate the measurement differences between the two software. Third, we did not evaluate the reproducibility of the software using a different MR scanner. Volumetric variability when using different MR scanners

may occur despite using the same volumetric software (32).

In conclusion, the SM and NQ provided more than moderate reliability for volumetry of most brain structures, except for the basal ganglia, in patients with cognitive impairment. However, volumetric estimates significantly differed for almost all brain structures, except the right HIP. The left HIP had minimal volume difference between the two software. Clinicians and researchers should be aware of these caveat when using these software in clinical practice.

Acknowledgments

REFERENCES

1. Park M, Moon WJ. Structural MR imaging in the diagnosis of Alzheimer's disease and other neurodegenerative dementia: current imaging approach and future perspectives. Korean J Radiol 2016;17:827-845
2. Chetelat G. Multimodal neuroimaging in Alzheimer's disease: early diagnosis, physiopathological mechanisms, and impact of lifestyle. J Alzheimers Dis 2018;64:S199-S211
3. Groot C, van Loenhoud AC, Barkhof F, et al. Differential effects of cognitive reserve and brain reserve on cognition in Alzheimer disease. Neurology 2018;90:e149-e156

4. Min J, Moon WJ, Jeon JY, Choi JW, Moon YS, Han SH. Diagnostic efficacy of structural MRI in patients with mild-to-moderate Alzheimer disease: automated volumetric assessment versus visual assessment. AJR Am J Roentgenol 2017;208:617-623

5. Ross DE, Ochs AL, DeSmit ME, Seabaugh JM, Havranek MD; Alzheimer's Disease Neuroimaging Initiative. Man versus machine Part 2: Comparison of radiologists' interpretations and NeuroQuant measures of brain asymmetry and progressive atrophy in patients with traumatic brain injury. J Neuropsychiatry Clin Neurosci 2015;27:147-152

6. Ross DE, Seabaugh J, Cooper L, Seabaugh J. NeuroQuant®and NeuroGage® reveal effects of traumatic brain injury on brain volume. Brain Inj 2018;32:1437-1441

7. Steenwijk MD, Amiri H, Schoonheim MM, et al. Agreement of MSmetrix with established methods for measuring cross-sectional and longitudinal brain atrophy. Neuroimage Clin 2017;15:843-853

8. Lee JS, Kim C, Shin JH, et al. Machine learning-based individual assessment of cortical atrophy pattern in Alzheimer's disease spectrum: development of the classifier and longitudinal evaluation. Sci Rep 2018;8:4161

9. Tanpitukpongse TP, Mazurowski MA, Ikhena J, Petrella JR; Alzheimer's Disease Neuroimaging Initiative. Predictive utility of marketed volumetric software tools in subjects at risk for Alzheimer disease: do regions outside the hippocampus matter? AJNR Am J Neuroradiol 2017;38:546-552

10. Persson K, Barca ML, Cavallin L, et al. Comparison of automated volumetry of the hippocampus using NeuroQuant® and visual assessment of the medial temporal lobe in Alzheimer's disease. Acta Radiol 2018;59:997-1001

11. Niemantsverdriet E, Ribbens A, Bastin C, et al. A Retrospective Belgian multi-center MRI biomarker study in Alzheimer's disease (REMEMBER). J Alzheimers Dis 2018;63:1509-1522

12. Ross DE, Ochs AL, Tate DF, et al. High correlations between MRI brain volume measurements based on NeuroQuant® and FreeSurfer. Psychiatry Res Neuroimaging 2018;278:69-76

13. Collins DL, Pruessner JC. Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. Neuroimage 2010;52:1355-1366

14. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999;56:303-308

15. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939-944

16. Roman GC, Tatemichi TK, Erkinjuntti T, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology 1993;43:250-260

17. Ryu HJ, Kim M, Moon Y, et al. Validation of the Korean version of the Lewy Body Composite Risk Score (K-LBCRS). J Alzheimers Dis 2017;55:1395-1401

18. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain 2011;134:2456-2477

19. Poewe W, Gauthier S, Aarsland D, et al. Diagnosis and management of Parkinson's disease dementia. Int J Clin Pract 2008;62:1581-1587

20. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155-163

21. Olejnik S, Algina J. Measures of effect size for comparative studies: applications, interpretations, and limitations. Contemp Educ Psychol 2000;25:241-286

22. Laubach M, Lammers F, Zacharias N, et al. Size matters: grey matter brain reserve predicts executive functioning in the elderly. Neuropsychologia 2018;119:172-181

23. Ochs AL, Ross DE, Zannoni MD, Abildskov TJ, Bigler ED; Alzheimer's Disease Neuroimaging Initiative. Comparison of automated brain volume measures obtained with NeuroQuant® and FreeSurfer. J Neuroimaging 2015;25:721-727

24. Reid MW, Hannemann NP, York GE, et al. Comparing two processing pipelines to measure subcortical and cortical volumes in patients with and without mild traumatic brain injury. J Neuroimaging 2017;27:365-371

25. Jack CR Jr, Therneau TM, Weigand SD, et al. Prevalence of biologically vs clinically defined alzheimer spectrum entities using the national institute on aging-Alzheimer's association research framework. JAMA Neurol 2019;76:1174-1183

26. Wang C, Beadnall HN, Hatton SN, et al. Automated brain volumetrics in multiple sclerosis: a step closer to clinical application. J Neurol Neurosurg Psychiatry 2016;87:754-757

27. Schmitter D, Roche A, Marechal B, et al. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. Neuroimage Clin 2015;7:7-17

28. Roche A, Marechal B, Kober T, et al. Assessing brain volumes using MorphoBox prototype. MAGNETOM Flash 2017;68:33-37

iMRI

Inter-Method Difference of Volumetric Softwares | Jieun Chung, et al.

29. Ogawa A, Yamazaki Y, Ueno K, Cheng K, Iriki A. Inferential reasoning by exclusion recruits parietal and prefrontal cortices. Neuroimage 2010;52:1603-1610

30. Haller S, Falkovskiy P, Meuli R, et al. Basic MR sequence parameters systematically bias automated brain volume estimation. Neuroradiology 2016;58:1153-1160

31. Stelmokas J, Yassay L, Giordani B, et al. Translational MRI volumetry with NeuroQuant: effects of version and normative data on relationships with memory performance in healthy older adults and patients with mild cognitive impairment. J Alzheimers Dis 2017;60:1499-1510

32. Guo C, Ferreira D, Fink K, Westman E, Granberg T. Repeatability and reproducibility of FreeSurfer, FSL-SIENAX and SPM brain volumetric measurements and the effect of lesion filling in multiple sclerosis. Eur Radiol 2019;29:1355-1364