



# Reliability of Changes in Brain Volume Determined by Longitudinal Voxel-Based Morphometry

Hidemasa Takao, MD, PhD,\*  Shiori Amemiya, MD, PhD,  and Osamu Abe, MD, PhD,,  
for the Alzheimer's Disease Neuroimaging Initiative<sup>†</sup>

**Background:** Longitudinal magnetic resonance imaging (MRI) studies have become increasingly important to assess the changes in brain morphology during normal aging and neurodegenerative disorders. However, the reliability of longitudinal morphometric changes has not been fully evaluated.

**Purpose:** To examine the reliability of longitudinal (2-year) changes in brain morphology determined by longitudinal voxel-based morphometry (VBM) in healthy elderly subjects, patients with mild cognitive impairment (MCI), and patients with Alzheimer's disease (AD).

**Study Type:** Retrospective analysis.

**Subjects:** Twenty-four healthy elderly subjects, 28 MCI patients, and 16 AD patients.

**Field Strength/Sequence:** A 1.5 T, magnetization-prepared rapid gradient-echo.

**Assessment:** Longitudinal (2-year) changes in gray matter volume determined by longitudinal VBM processing, and visual assessment of image quality.

**Statistical Tests:** Intraclass correlation coefficient (ICC) and Kruskal–Wallis test.

**Results:** The ICC maps differed among the three groups. The mean ICC was 0.81 overall (0.86 for healthy elderly subjects, 0.75 for MCI patients, and 0.76 for AD patients). The reliability was good to excellent (ICC, 0.60–1.00) for 92% of voxels (99% for healthy elderly subjects, 83% for MCI patients, and 83% for AD patients). The image quality differed significantly among the three groups ( $P < 0.05$ ).

**Data Conclusion:** These results indicate that the reliability of longitudinal gray matter volume changes by VBM is good to excellent for most voxels. However, reliability may be affected by the disease, possibly due to differences in head motion during imaging.

**Evidence Level:** 3

**Technical Efficacy:** Stage 1

J. MAGN. RESON. IMAGING 2021.

## Introduction

Knowledge about the effects of normal aging and neurodegenerative disorders on brain morphology is mainly derived from cross-sectional studies; however, extensive between-subject variability in brain morphology reduces the sensitivity for detecting changes in brain morphology.<sup>1</sup> Within-subject changes in morphology are usually smaller than between-subject differences.<sup>1</sup> Longitudinal studies, which can avoid

some of the problems of secular trends and between-subject variation, have become increasingly important in assessing the changes in brain morphology that occur during normal aging and neurodegenerative disorders.<sup>2</sup> Longitudinal studies limit the variability in brain morphology attributable to between-subject differences by using each subject as their own control.<sup>1</sup> Despite this, the statistical power to detect changes in brain morphology can be limited by measurement errors.<sup>1</sup> To

View this article online at [wileyonlinelibrary.com](http://wileyonlinelibrary.com). DOI: 10.1002/jmri.27568

Received Dec 1, 2020, Accepted for publication Feb 3, 2021.

\*Address reprint requests to: H. T., Department of Radiology, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. E-mail: [takaoh-ky@umin.ac.jp](mailto:takaoh-ky@umin.ac.jp)

<sup>†</sup>Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

From the Department of Radiology, Graduate School of Medicine, University of Tokyo, Tokyo, 113-8655, Japan

quantify the rates of changes in brain morphology from serial MRI scans, it is important that the acquisitions at baseline and at a later time are as similar as possible.<sup>3</sup>

Sufficient reliability is essential to use neuroimaging as a potential biomarker of neurodegenerative disorders, especially when monitoring the longitudinal changes associated with a disorder and the effects of treatment.<sup>3</sup> Although previous studies have evaluated the reliability of structural  $T_1$ -weighted imaging and diffusion imaging,<sup>1,4–24</sup> the reliability of the estimated longitudinal changes in brain morphometry has not been fully evaluated, and it is unclear whether neurological disorders affect this reliability.

The purpose of this study was to examine the reliability of longitudinal (2-year) changes in brain morphology determined by longitudinal voxel-based morphometry (VBM) among healthy elderly subjects, patients with mild cognitive impairment (MCI), and patients with Alzheimer's disease (AD).

## Materials and Methods

### Subjects

This study used data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (available at <http://adni.loni.usc.edu>). The ADNI was launched in 2003 as a public-private partnership, led by the Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI was to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. The ADNI was approved by the institutional review boards of all participating sites. Written informed consent was obtained from all participants.

This study used the ADNI 1 database to select subjects on whom two back-to-back, structural  $T_1$ -weighted imaging scans, on 1.5 T Philips scanners, had been performed twice (at screening and a 2-year follow-up; two scans  $\times$  two time-points per subject). Subjects were selected by the following criteria: availability of two back-to-back, structural  $T_1$ -weighted imaging data, on 1.5 T Philips scanners, at screening and 2-year follow-up. Subjects who had a lack of imaging data were excluded. A total of 68 subjects (24 healthy control subjects, 28 patients with MCI, and 16 patients with AD) were included in this study. The mean age (range) at screening was  $75 \pm 7$  years (healthy control subjects,  $75 \pm 4$  years [70–86 years]; patients with MCI,  $76 \pm 7$  years [56–87 years]; and patients with AD,  $73 \pm 8$  years [60–85 years]). The mean scan interval (range) was  $2.1 \pm 0.1$  years (healthy control subjects,  $2.1 \pm 0.1$  years [2.0–2.4 years]; patients with MCI,  $2.1 \pm 0.1$  years [2.0–2.5 years]; and patients with AD,  $2.1 \pm 0.1$  years [2.0–2.4 years]).

### Imaging Data Acquisition

MRI was performed using 1.5 T Philips scanners at multiple sites using the ADNI 1.5 T imaging protocol. Various models of scanners were used in the ADNI (for details, please refer to <http://adni.loni.usc.edu>), but each subject underwent scans at screening and follow-up on the same scanner. Structural  $T_1$ -weighted images were acquired using a three-dimensional magnetization-prepared rapid gradient-echo (MP-

RAGE) sequence in 170 sagittal slices (repetition time = 8.6 msec; echo time = 4.0 msec; inversion time = 1000 msec; flip angle =  $8^\circ$ ; field of view =  $240 \times 240$  mm; slice thickness = 1.2 mm with no gap; acquisition matrix =  $192 \times 192$ ; image matrix =  $256 \times 256$ ; reconstructed voxel size =  $0.94 \times 0.94 \times 1.2$  mm). The non-uniform intensity of MP-RAGE images was corrected using the nonparametric non-uniform intensity normalization algorithm N3.<sup>25–27</sup>

### Image Quality

The quality of MP-RAGE images was subjectively graded as good, adequate, or poor upon visual inspection. Three radiologists with 21 (H.T.), 10, and 2 years of experience in neuroradiology independently evaluated the image quality in a blinded manner. In case of disagreements, final evaluations were made by consensus. The MP-RAGE images obtained at screening and the 2-year follow-up were divided into pairs of higher-quality and lower-quality images for subsequent analysis.

### Image Processing

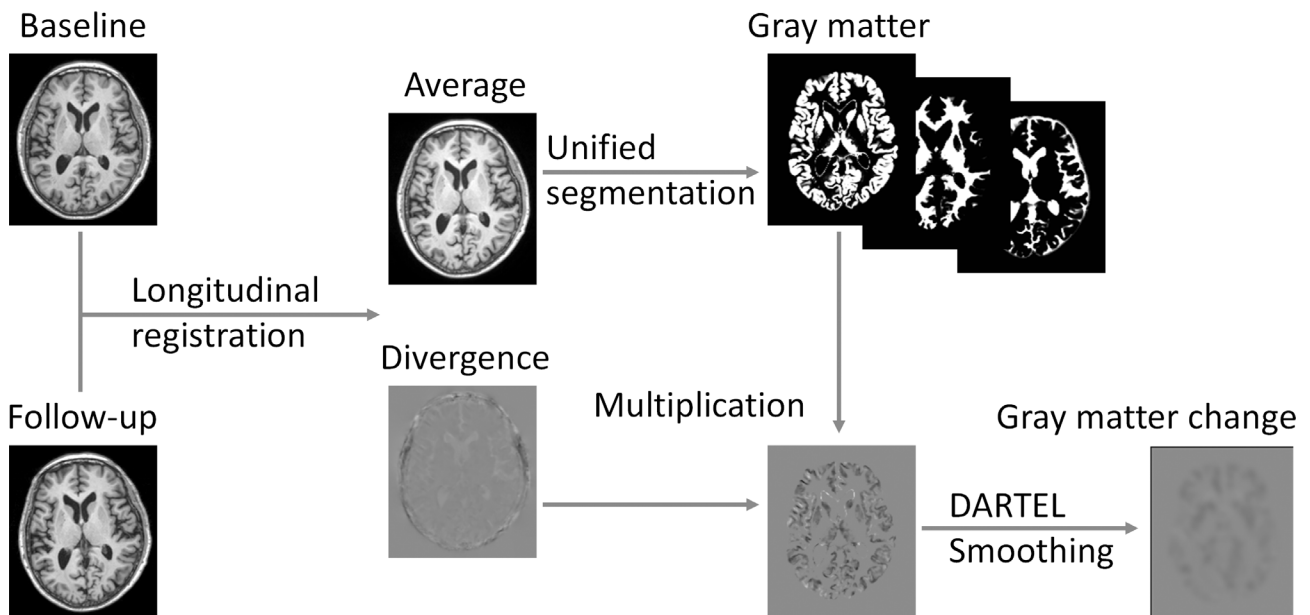
Images were mainly processed using statistical parametric mapping (SPM) 12 software (<http://www.fil.ion.ucl.ac.uk/spm>) developed by the Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College London, running in MATLAB 9.1 (Mathworks, Sherborn, MA) (Figure 1).

The longitudinal registration of pairs of MP-RAGE images (obtained at screening and 2-year follow-up) was performed using pairwise inverse-consistent alignment between the first and second scan of each subject, incorporating a bias field correction to calculate the mid-point average image and a map of the divergence of the velocity fields (representing the rates of volumetric expansion/contraction) for each subject.<sup>28</sup> The mid-point average images were segmented into gray matter, white matter, and cerebrospinal fluid using an integrated generative model (unified segmentation).<sup>29</sup> The International Consortium for Brain Mapping (ICBM) gray matter, white matter, cerebrospinal fluid, bone, soft tissue, and air/background templates were used as priors to segment the images.

The Diffeomorphic Anatomical Registration through Exponentiated Lie Algebra (DARTEL) algorithm was used to spatially normalize the segmented gray matter and white matter images, and the images of longitudinal gray matter volume changes calculated by multiplying the gray matter images by the divergence maps.<sup>30</sup> The normalized images were modulated to correct the voxel signal intensity for volume displacement during normalization and hence reflect brain volume. The images were smoothed using an 8-mm kernel.

### Statistical Analysis

The intraclass correlation coefficient (ICC) was calculated for each voxel in the images of longitudinal gray matter volume changes based on a single-measurement, absolute-agreement, two-way mixed-effects model using MATLAB 9.1.<sup>31,32</sup> Histogram analysis was performed for each ICC map with a histogram bin width of 0.002 and a range of  $-1.0$  to  $1.0$ . Only voxels with a volume  $> 0.05$  (on all gray matter images) were included in the ICC calculation and histogram analysis. The ICC was interpreted using the criteria proposed by Cicchetti et al,<sup>33</sup> where an ICC of  $< 0.40$  is considered poor,  $0.40$ – $0.59$  is fair,  $0.60$ – $0.74$  is good, and  $0.75$ – $1.00$  is excellent.



**FIGURE 1: Summary of longitudinal voxel-based morphometry (VBM) processing using statistical parametric mapping (SPM) 12 software. DARTEL, Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra.**

The mean and SD images of longitudinal gray matter volume changes were calculated from the pairs of high-quality images. The distribution of gray matter atrophy was visually assessed by a radiologist with 21 years of experience in neuroradiology (H.T.), and confirmed by two radiologists with 10 and 2 years of experience in neuroradiology.

To evaluate the effect of image quality on the reliability of longitudinal volume changes, the Kruskal–Wallis test was used to compare image quality among healthy control subjects, patients with MCI, and patients with AD using SPSS Statistics version 22 (IBM, Armonk, NY). The significance level was set at  $P < 0.05$ . The inter-rater reliability of image quality ratings was assessed using the ICC calculated based on a single-measurement, absolute-agreement, and two-way mixed-effects model.

## Results

### ICC Maps and Histograms

Figure 2 shows the voxel-wise ICC maps of the longitudinal (2-year) changes in gray matter volume. Figure 3 shows the results of the histogram analysis (frequency polygons) of the ICC maps. The ICC maps and their frequency polygons differed among the healthy control subjects, patients with MCI, and patients with AD. The mean ICC was 0.81 overall (0.86 for healthy control subjects, 0.75 for patients with MCI, and 0.76 for patients with AD). The histogram peak was 0.90 overall (0.93 for healthy control subjects, 0.90 for patients with MCI, and 0.95 for patients with AD). Table 1 summarizes the distribution of the voxel-wise ICC estimates. Overall, the reliability was excellent (ICC, 0.75–1.00) for 76% of voxels (91% for healthy control subjects, 63% for patients with MCI, and 64% for patients with AD). The reliability was good to excellent (ICC, 0.60–1.00) for 92% of voxels

(99% for healthy control subjects, 83% for patients with MCI, and 83% for patients with AD).

### Two-Year Changes in Gray Matter Volume

Figure 4 shows the mean longitudinal (2-year) changes in gray matter volume for healthy control subjects, patients with MCI, and patients with AD. Gray matter atrophy was more extensive in patients with MCI or AD than in healthy control subjects. Gray matter atrophy was particularly prominent in the temporal lobe, including the hippocampus and parahippocampal cortex, as well as the posterior cingulate cortex and precuneus. Figure 5 shows the standard deviations of longitudinal (2-year) changes in gray matter volume for healthy control subjects, patients with MCI, and patients with AD. As a whole, the variance of longitudinal volume changes was larger in patients with MCI or AD than in healthy control subjects.

### Image Quality

Figure 6 shows the distribution of image quality for healthy control subjects, patients with MCI, and patients with AD. The distribution of image quality (rated as good, adequate, or poor) differed significantly among the three groups (Kruskal–Wallis test,  $P < 0.05$ ). The inter-rater reliability of image quality ratings was good (ICC, 0.72).

### Discussion

The reliability of the longitudinal (2-year) changes in gray matter volume was good to excellent for most voxels. However, the reliability was worse in patients with MCI and patients with AD than in healthy elderly subjects. The image quality was significantly different among groups, with worse

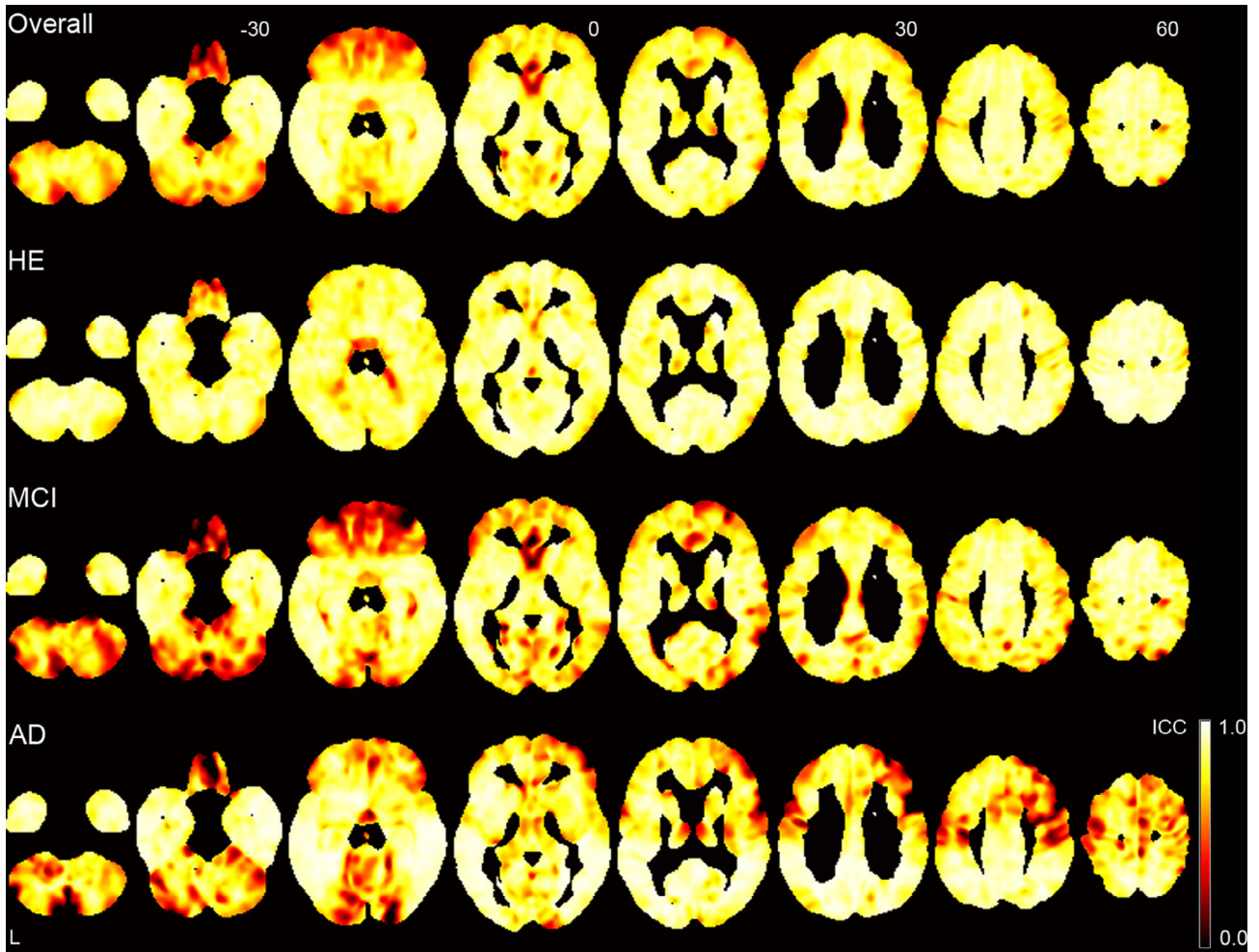


FIGURE 2: Voxel-wise intraclass correlation coefficient (ICC) maps of longitudinal (2-year) changes in gray matter volume. HE: healthy elderly; MCI: mild cognitive impairment; AD: Alzheimer’s disease.

quality in patients with MCI or AD than in healthy elderly subjects; the differences might be due to head motion. These findings might explain the differences in the reliability of the longitudinal (2-year) changes in gray matter volume among

healthy elderly subjects, patients with MCI, and patients with AD.

This study used the longitudinal registration implemented in SPM software to register baseline and follow-up scans and calculate longitudinal volume changes.<sup>28</sup> This approach combines rigid alignment, diffeomorphic warping, and differential intensity non-uniformity correction on a within-subject template that evolves to yield an average of all three of these aspects, and shows a symmetric, transitive construction. Longitudinal studies of brain morphology generally use longitudinal image processing, which reduces within-subject variability by calculating the within-subject changes integrating the information of scans at each time-point<sup>28,34</sup>; this differs from cross-sectional studies, which treat each scan at each time-point independently. Bias, however, can be introduced by longitudinal image processing when scans from different time-points are not treated equivalently and symmetrically as they undergo different processing steps. For example, interpolation asymmetries occur when registering follow-up scans to the baseline scan, while smoothing the follow-up scans and preserving the unaltered baseline scan.<sup>35</sup>

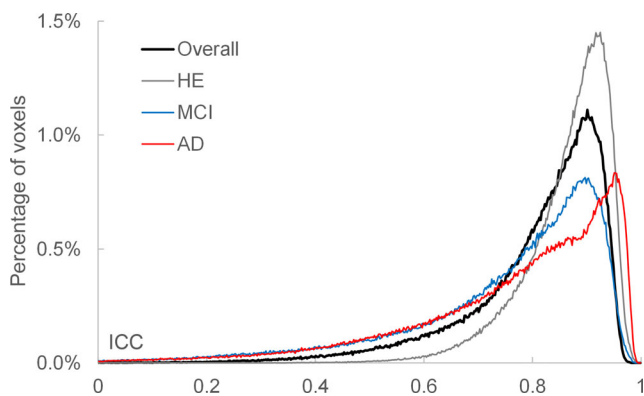
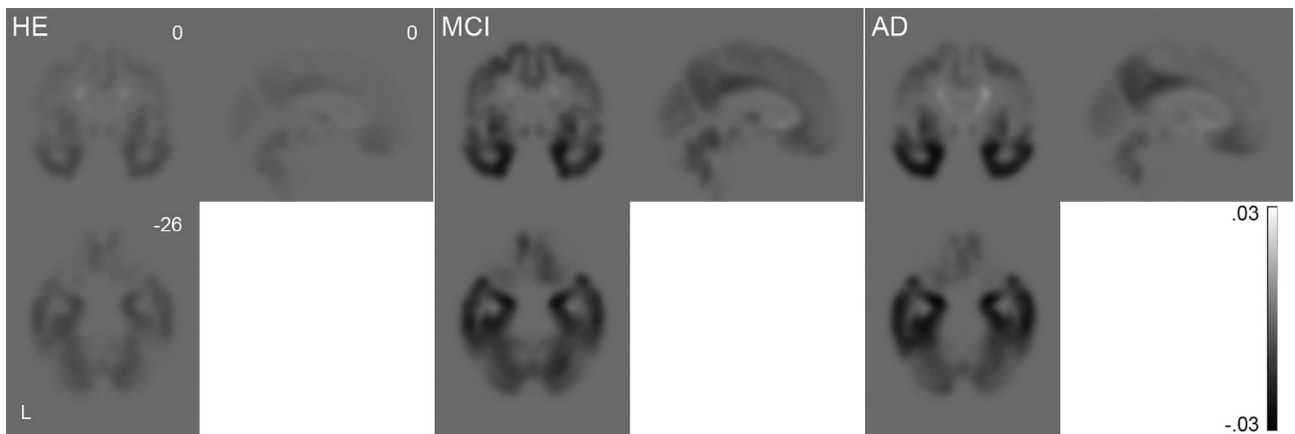


FIGURE 3: Histograms (frequency polygons) of voxel-wise intraclass correlation coefficient (ICC) maps of longitudinal (2-year) changes in gray matter volume. HE: healthy elderly; MCI: mild cognitive impairment; AD: Alzheimer’s disease.

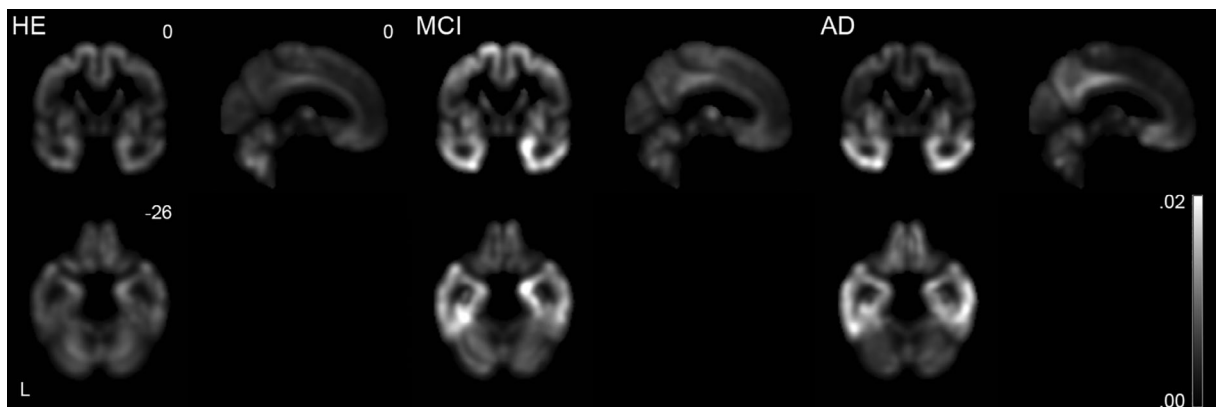
**TABLE 1. Distribution of voxel-wise intraclass correlation coefficient (ICC) estimates of longitudinal (2-year) changes in gray matter volume**

ICC	<0.40 (poor) (%)	0.40–0.59 (fair) (%)	0.60–0.74 (good) (%)	0.75–1.00 (excellent) (%)
Overall	2	6	16	76
HE	0	1	8	91
MCI	6	11	20	63
AD	6	11	19	64

HE: healthy elderly; MCI: mild cognitive impairment; AD: Alzheimer's disease.



**FIGURE 4: Mean longitudinal (2-year) changes in gray matter volume. HE: healthy elderly; MCI: mild cognitive impairment; AD: Alzheimer's disease.**

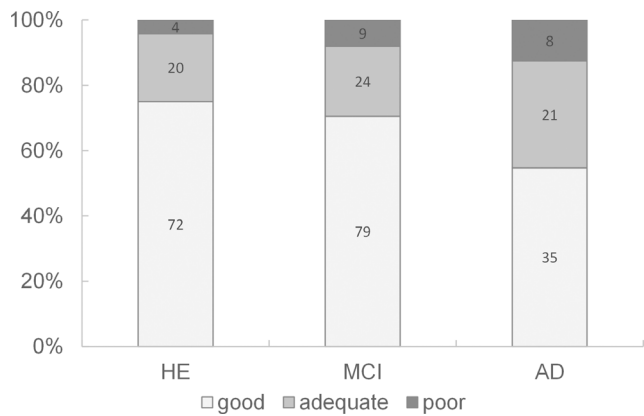


**FIGURE 5: Standard deviations of longitudinal (2-year) changes in gray matter volume. HE: healthy elderly; MCI: mild cognitive impairment; AD: Alzheimer's disease.**

To prevent bias from affecting the estimated longitudinal changes in brain morphology, it is essential to treat scans obtained at different times symmetrically; otherwise, longitudinal image processing can introduce more problems than it solves.

AD is the most common cause of dementia in older adults.<sup>36</sup> It is an irreversible, progressive neurodegenerative

disorder, in which amyloid plaques and neurofibrillary tangles accumulate in the brain, impairing axons, dendrites, and synapses.<sup>37</sup> Plaques are abnormal, dense deposits of amyloid  $\beta$  peptides that accumulate between neurons and disrupt cell function. Neurofibrillary tangles are abnormal aggregates of hyperphosphorylated tau protein that accumulate inside neurons, block neuronal transport systems, and affect neuronal



**FIGURE 6: Distribution of image quality.** HE: healthy elderly; MCI: mild cognitive impairment; AD: Alzheimer's disease.

synaptic communication.<sup>38</sup> AD is characterized by behavioral changes, memory loss, and deteriorating learning capacity, resulting in cognitive and behavioral impairments that affect and limit normal daily activities.<sup>37</sup> Cortical atrophy occurs earliest in the medial temporal lobe (entorhinal cortex and hippocampus) and subsequently extends to the remaining cortex along a temporal–parietal–frontal trajectory, while sensorimotor and visual cortices are generally spared until later stages.<sup>39</sup> This topographical progression is correlated with disease severity and the appearance of clinical symptoms.<sup>39</sup> In early stages, marked cortical atrophy occurs in the medial temporal lobe and posterior cingulate/retrosplenial cortex, and is accompanied by milder atrophy in the orbitofrontal cortex.<sup>39,40</sup> The subsequent progression from incipient to mild AD is associated with widespread and marked cortical atrophy, affecting broad areas of the lateral temporal cortex, dorsal parietal, and frontal cortex.<sup>39,40</sup> In the present study, longitudinal VBM detected longitudinal gray matter atrophy in similar regions.

This study found that the reliability of the longitudinal (2-year) changes in gray matter volume differed among healthy elderly subjects, patients with MCI, and patients with AD. Head motion is a key contributor to the within-subject variability in a variety of neuroimaging modalities.<sup>41,42</sup> Although head motion cannot be directly measured during structural T<sub>1</sub>-weighted imaging, it impacts on image quality.<sup>42</sup> In the present study using ADNI data, head motion could not be directly measured and instead evaluated in terms of image quality. The image quality was significantly different among healthy elderly subjects, patients with MCI, and patients with AD, and was worse in patients with MCI or AD than in healthy elderly subjects. These results suggest that head motion may influence the reliability of the longitudinal changes in gray matter volume in healthy elderly subjects, patients with MCI, and patients with AD.

In this study, the variance of longitudinal (2-year) changes in gray matter volume was overall larger in patients

with MCI or AD than in healthy elderly subjects. By definition, a larger variance between subjects compared with measurement error results in larger ICC. Therefore, the larger variance of longitudinal volume changes could not explain the lower ICC in patients with MCI and AD.

## Limitations

This study has some limitations. First, the condition of the subjects might change during the 2-year follow-up period. Longitudinal studies of brain morphometry usually have follow-up periods of 1–2 or more years. To evaluate the reliability of longitudinal morphometric changes determined by longitudinal voxel-based morphometry under realistic conditions, we decided to use scans obtained at the 2-year interval. Second, head motion was evaluated in terms of image quality. This may have a limited ability to measure head motion compared with direct measurement. Finally, the effect of site/scanner on longitudinal morphometric changes was not evaluated. Although each subject underwent scans at screening and follow-up on the same scanner, the ADNI used various models of scanners at various sites. The effect of site/scanner on longitudinal morphometric changes might exist but is somewhat beyond the scope of this study.

## Conclusion

The results of this study indicate that the reliability of longitudinal changes in gray matter volume by VBM are good to excellent in most voxels, but the reliability may be affected by the underlying disease. These differences may be due to differences in head motion during imaging among healthy elderly subjects, patients with MCI, and patients with AD.

## Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale



Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (<https://www.fnih.org>). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

## References

1. Takao H, Hayashi N, Ohtomo K. Effects of the use of multiple scanners and of scanner upgrade in longitudinal voxel-based morphometry studies. *J Magn Reson Imaging* 2013;38:1283-1291.
2. Takao H, Hayashi N, Ohtomo K. A longitudinal study of brain volume changes in normal aging. *Eur J Radiol* 2012;81:2801-2804.
3. Jack CR Jr, Bernstein MA, Fox NC, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging* 2008;27:685-691.
4. Schnack HG, van Haren NE, Hulshoff Pol HE, et al. Reliability of brain volumes from multicenter MRI acquisition: A calibration study. *Hum Brain Mapp* 2004;22:312-320.
5. Ewers M, Teipel SJ, Dietrich O, et al. Multicenter assessment of reliability of cranial MRI. *Neurobiol Aging* 2006;27:1051-1059.
6. Han X, Jovicich J, Salat D, et al. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *Neuroimage* 2006;32:180-194.
7. Dickerson BC, Fenstermacher E, Salat DH, et al. Detection of cortical thickness correlates of cognitive performance: Reliability across MRI scan sessions, scanners, and field strengths. *Neuroimage* 2008;39:10-18.
8. Jovicich J, Czanner S, Han X, et al. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *Neuroimage* 2009;46:177-192.
9. Huppertz HJ, Kroll-Seeger J, Kloppel S, Ganz RE, Kassubek J. Intra- and interscanner variability of automated voxel-based volumetry based on a 3D probabilistic atlas of human cerebral structures. *Neuroimage* 2010;49:2216-2224.
10. Takao H, Hayashi N, Ohtomo K. Effect of scanner in longitudinal studies of brain volume changes. *J Magn Reson Imaging* 2011;34:438-444.
11. Cannon TD, Sun F, McEwen SJ, et al. Reliability of neuroanatomical measurements in a multisite longitudinal study of youth at risk for psychosis. *Hum Brain Mapp* 2014;35:2424-2434.
12. Takao H, Hayashi N, Ohtomo K. Effects of study design in multi-scanner voxel-based morphometry studies. *Neuroimage* 2014;84:133-140.
13. Takao H, Hayashi N, Ohtomo K. Brain morphology is individual-specific information. *Magn Reson Imaging* 2015;33:816-821.
14. Biberacher V, Schmidt P, Keshavan A, et al. Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. *Neuroimage* 2016;142:188-197.
15. Lee H, Nakamura K, Narayanan S, Brown RA, Arnold DL. Estimating and accounting for the effect of MRI scanner changes on longitudinal whole-brain volume change measurements. *Neuroimage* 2019;184:555-565.
16. Melzer TR, Keenan RJ, Leeper GJ, et al. Test-retest reliability and sample size estimates after MRI scanner relocation. *Neuroimage* 2020;211:116608.
17. Heiervang E, Behrens TE, Mackay CE, Robson MD, Johansen-Berg H. Between session reproducibility and between subject variability of diffusion MR and tractography measures. *Neuroimage* 2006;33:867-877.
18. Vollmar C, O'Muircheartaigh J, Barker GJ, et al. Identical, but not the same: Intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0T scanners. *Neuroimage* 2010;51:1384-1394.
19. Takao H, Hayashi N, Ohtomo K. Effect of scanner in asymmetry studies using diffusion tensor imaging. *Neuroimage* 2011;54:1053-1062.
20. Zhu T, Hu R, Qiu X, et al. Quantification of accuracy and precision of multi-center DTI measurements: A diffusion phantom and human brain study. *Neuroimage* 2011;56:1398-1411.
21. Lemkaddem A, Daducci A, Vulliemoz S, et al. A multi-center study: Intra-scan and inter-scan variability of diffusion spectrum imaging. *Neuroimage* 2012;62:87-94.
22. Takao H, Hayashi N, Kabasawa H, Ohtomo K. Effect of scanner in longitudinal diffusion tensor imaging studies. *Hum Brain Mapp* 2012;33:466-477.
23. Wang JY, Abdi H, Bakhadirov K, Diaz-Arrastia R, Devous MD Sr. A comprehensive reliability assessment of quantitative diffusion tensor tractography. *Neuroimage* 2012;60:1127-1138.
24. Takao H, Hayashi N, Ohtomo K. Brain diffusivity pattern is individual-specific information. *Neuroscience* 2015;301:395-402.
25. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 1998;17:87-97.
26. Takao H, Abe O, Hayashi N, Kabasawa H, Ohtomo K. Effects of gradient non-linearity correction and intensity non-uniformity correction in longitudinal studies using structural image evaluation using normalization of atrophy (SIENA). *J Magn Reson Imaging* 2010;32:489-492.
27. Takao H, Abe O, Ohtomo K. Computational analysis of cerebral cortex. *Neuroradiology* 2010;52:691-698.
28. Ashburner J, Ridgway GR. Symmetric diffeomorphic modeling of longitudinal structural MRI. *Front Neurosci* 2013;6:197-197.
29. Ashburner J, Friston KJ. Unified segmentation. *Neuroimage* 2005;26:839-851.
30. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage* 2007;38:95-113.
31. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155-163.
32. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30-46.
33. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess* 1994;6:284-290.
34. Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* 2012;61:1402-1418.
35. Reuter M, Fischl B. Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage* 2011;57:19-21.
36. Mangialasche F, Solomon A, Winblad B, Mecocci P, Kivipelto M. Alzheimer's disease: Clinical trials and drug development. *Lancet Neurol* 2010;9:702-716.
37. Pietzuch M, King AE, Ward DD, Vickers JC. The influence of genetic factors and cognitive reserve on structural and functional resting-state brain networks in aging and Alzheimer's disease. *Front Aging Neurosci* 2019;11:30.

38. Spiess-Jones TL, Hyman BT. The intersection of amyloid beta and tau at synapses in Alzheimer's disease. *Neuron* 2014;82:756-771.
39. Pini L, Pievani M, Bocchetta M, et al. Brain atrophy in Alzheimer's disease and aging. *Ageing Res Rev* 2016;30:25-48.
40. Frisoni GB, Prestia A, Rasser PE, Bonetti M, Thompson PM. In vivo mapping of incremental cortical atrophy from incipient to overt Alzheimer's disease. *J Neurol* 2009;256:916-924.
41. Zuo XN, Xing XX. Test-retest reliabilities of resting-state fMRI measurements in human brain functional connectomics: A systems neuroscience perspective. *Neurosci Biobehav Rev* 2014;45:100-118.
42. Savalia NK, Agres PF, Chan MY, Feczko EJ, Kennedy KM, Wig GS. Motion-related artifacts in structural brain images revealed with independent estimates of in-scanner head motion. *Hum Brain Mapp* 2017;38:472-492.