AJNR

Evaluation of Ultrafast Wave-CAIPI MPRAGE for Visual Grading and Automated Measurement of Brain Tissue Volume

M.G.F. Longo, J. Conklin, S.F. Cauley, K. Setsompop, Q. Tian, D. Polak, M. Polackal, D. Splitthoff, W. Liu, R.G. González, P.W. Schaefer, J.E. Kirsch, O. Rapalino and S.Y. Huang

AJNR Am J Neuroradiol 2020, 41 (8) 1388-1396
doi: https://doi.org/10.3174/ajnr.A6703
http://www.ajnr.org/content/41/8/1388

Evaluation of Ultrafast Wave-CAIPI MPRAGE for Visual Grading and Automated Measurement of Brain Tissue Volume

M.G.F. Longo, J. Conklin, S.F. Cauley, K. Setsompop, Q. Tian, D. Polak, M. Polackal, D. Splitthoff, W. Liu, R.G. González, P.W. Schaefer, J.E. Kirsch, O. Rapalino, and S.Y. Huang

ABSTRACT

BACKGROUND AND PURPOSE: Volumetric brain MR imaging typically has long acquisition times. We sought to evaluate an ultrafast MPRAGE sequence based on Wave-CAIPI (Wave-MPRAGE) compared with standard MPRAGE for evaluation of regional brain tissue volumes.

MATERIALS AND METHODS: We performed scan-rescan experiments in 10 healthy volunteers to evaluate the intraindividual variability of the brain volumes measured using the standard and Wave-MPRAGE sequences. We then evaluated 43 consecutive patients undergoing brain MR imaging. Patients underwent 3T brain MR imaging, including a standard MPRAGE sequence (acceleration factor [*R*] = 2, acquisition time [TA] = 5.2 minutes) and an ultrafast Wave-MPRAGE sequence (*R* = 9, TA = 1.15 minutes for the 32-channel coil; *R* = 6, TA = 1.75 minutes for the 20-channel coil). Automated segmentation of regional brain volume was performed. Two radiologists evaluated regional brain atrophy using semiquantitative visual rating scales.

RESULTS: The mean absolute symmetrized percent change in the healthy volunteers participating in the scan-rescan experiments was not statistically different in any brain region for both the standard and Wave-MPRAGE sequences. In the patients undergoing evaluation for neurodegenerative disease, the Dice coefficient of similarity between volumetric measurements obtained from standard and Wave-MPRAGE ranged from 0.86 to 0.95. Similarly, for all regions, the absolute symmetrized percent change for brain volume and cortical thickness showed <6% difference between the 2 sequences. In the semiquantitative visual comparison, the differences between the 2 radiologists' scores were not clinically or statistically significant.

CONCLUSIONS: Brain volumes estimated using ultrafast Wave-MPRAGE show low intraindividual variability and are comparable with those estimated using standard MPRAGE in patients undergoing clinical evaluation for suspected neurodegenerative disease.

ABBREVIATIONS: ASPC = absolute symmetrized percent change; VBM = voxel-based morphometry

Volumetric brain MR imaging is widely used in clinical and research settings for the evaluation of patients with suspected neurodegenerative disease. Regional patterns of tissue loss aid in generating a differential diagnosis and assessing prognosis, and the identification of regional volume loss is increasingly used as an outcome measure in trials of potentially disease-modifying therapies.[1-4] Of particular value, the T1-weighted MPRAGE sequence provides excellent spatial resolution and tissue contrast[5] but has long acquisition times due to the need to encode a large number of *k*-space lines and the added TI required to achieve the prepared T1-weighted contrast. Unfortunately, long

Received January 25, 2020; accepted after revision May 18.

From the Departments of Radiology (M.G.F.L., J.C., M.P., R.G.G., P.W.S., J.E.K., O.R., S.Y.H.) and Radiology, Athinoula A. Martinos Center for Biomedical Imaging, (J.C., S.F.C., K.S., Q.T., D.P., S.Y.H.), Massachusetts General Hospital, Boston, Massachusetts; Harvard Medical School (J.C., S.F.C., K.S., R.G.G., P.W.S., S.Y.H.), Boston, Massachusetts; Harvard-MIT Division of Health Sciences and Technology (K.S., S.Y.H.), Massachusetts Institute of Technology, Cambridge, Massachusetts; Department of Physics and Astronomy (D.P.), Heidelberg University, Heidelberg, Germany; and Siemens (D.P., D.S., W.L.), Erlangen, Germany

M.G.F. Longo and J. Conklin are co-first authors.

The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic health care centers, or the National Institutes of Health.

This work was supported by the National Institutes of Health (P41 EB015896, R01 EB020613, UL 1TR002541), a Radiological Society of North America Research Resident Grant (S.Y.H.), and Siemens. It was conducted with support from Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Advancing Translational Sciences, National Institutes of Health Award UL 1TR002541) and financial contributions from Harvard University and its affiliated academic health care centers.

Please address correspondence to Maria G.F. Longo, MD, MSc, Division of Emergency Radiology, Department of Radiology, Massachusetts General Hospital, 55 Fruit St, Blk SB Rm 0029A, Boston, MA, 02114; e-mail: mfigueirolongo@mgh.harvard.edu

Indicates open access to non-subscribers at www.ajnr.org
Indicates article with supplemental on-line tables.
Indicates article with supplemental on-line photos.
http://dx.doi.org/10.3174/ajnr.A6703

Table 1: Acquisition parameters for standard and wave MPRAGE sequences

Parameters	Standard MPRAGE	Wave MPRAGE
FOV read (mm)	240 × 240	256 × 256
FOV phase (%)	100	100
Matrix size	256 × 256	256 × 256
Section thickness (mm)	0.89	1.0
TR/TE/TI (ms)	2300/2.32/900	2500/3.48/1100
Flip angle	8°	7°
Acceleration factor		
20-Channel	GRAPPA, $R = 2$	Wave-CAIPI, $R = 3 \times 2$
32-Channel	GRAPPA, $R = 2$	Wave-CAIPI, $R = 3 \times 3$
Bandwidth (Hx/px)	200	200
Scan time (sec)		
20-Channel	5 min 19 sec	1 min 46 sec
32-Channel	5 min 18 sec	1 min 11 sec

Note:—GRAPPA indicates generalized autocalibrating partially parallel acquisition.

scan times can contribute to patient anxiety, motion artifacts, and nondiagnostic examinations, particularly for motion-prone elderly patients.[6-8]

The Wave-Controlled Aliasing in Parallel Imaging (CAIPI; Siemens) acquisition and reconstruction approach enables up to an order of magnitude reduction in scan time with relatively preserved image quality and clinically feasible reconstruction times.[9,10] The adoption of Wave-CAIPI technology for highly accelerated imaging in clinical and research studies will be facilitated by systematic validation of its use in routine clinical imaging protocols.[11] Wave-CAIPI has been optimized for whole-brain imaging with MPRAGE in healthy volunteers[12] and has demonstrated potential in accelerating whole-brain volumetric evaluation of healthy volunteers[13] but has not yet been systematically evaluated in a clinical setting.

The purpose of this study was to compare a highly accelerated MPRAGE acquisition based on Wave-CAIPI (Wave-MPRAGE) with standard MPRAGE in a prospective study of patients undergoing evaluation for suspected neurodegenerative disease. This evaluation consisted of the following: 1) assessment of scan-rescan reliability of brain volumes extracted using each sequence in a small cohort of healthy volunteers; 2) comparison of the 2 sequences using automated measures of regional brain volume and cortical thickness in a larger cohort of patients undergoing evaluation for neurodegenerative disease; and 3) comparison of the 2 sequences using 6 histopathologically validated visual rating scales of regional brain atrophy. We hypothesized that Wave-MPRAGE would be the equivalent to standard MPRAGE for these quantitative and semiquantitative evaluations of regional brain volume, despite a 3- to 4-fold decrease in acquisition time.

MATERIALS AND METHODS

Subjects and Study Design

All components of this study were Health Insurance Portability and Accountability Act–compliant and underwent approval by the institutional review board (Massachusetts General Hospital). We split our study into 2 parts. We first performed scan-rescan experiments using the standard and Wave-MPRAGE sequences in a small cohort of healthy volunteers. We then compared the standard and Wave-MPRAGE images in a larger cohort of patients undergoing evaluation for neurodegenerative disease.

Scan-Rescan Experiments in Healthy Volunteers. Ten healthy volunteers were recruited for scan-rescan experiments to assess the scan-rescan reliability of brain volumes extracted using each sequence. Written informed consent was obtained from all subjects before MR imaging scanning. All images were acquired in the same 3T MR imaging scanner (Magnetom Skyra; Siemens). Twenty- or 32-channel multiarray receiver coils were used, with 5 subjects randomly chosen to be scanned using each head coil. The scan sessions were divided into 2 parts. Wave and standard MPRAGE sequences (sequence parameters provided in Table 1) were acquired in a randomized order during each part of the scan session. The subjects were removed from the scanner and repositioned between the first and second half of their scan sessions.

Wave-versus-Standard MPRAGE Evaluation in Neurodegenerative Disease. A prospective comparative study of Wave versus standard MPRAGE was performed at a single institution from July 2018 to February 2019. Patients undergoing brain MR imaging as part of the clinical work-up for memory loss or suspected neurodegenerative disease ($n = 31$) or as part of a concurrent research study ($n = 12$) were prospectively enrolled. From the patients participating in the research studies, 8 (66.7%) were part of a study investigating chronic white matter disease, 3 (25.0%) were part of a study investigating ischemic stroke, and 1 (8.3%) was part of an epilepsy study. There were no exclusion criteria beyond those for routine clinical MR imaging. For the 31 subjects undergoing MR imaging as part of their clinical evaluation, verbal consent was obtained at the time of MR imaging and written consent was waived by the institutional review board. For the 12 subjects undergoing MR imaging solely for research purposes, written informed consent was obtained before MR imaging.

Wave-MPRAGE Sequence

Wave-MPRAGE was implemented using a prototype inversion-recovery prepared 3D gradient-echo pulse sequence[12] (Wave-CAIPI; Siemens Healthineers, Erlangen, Germany). On-line reconstruction was performed using an auto-calibrated procedure for simultaneous estimation of the parallel imaging reconstruction and true *k*-space trajectory,[10] with an on-line reconstruction time of approximately 60 seconds. Pulse sequence parameters could not be exactly matched between the Wave-MPRAGE and standard MPRAGE sequences due to vendor constraints on the available parameter options (eg, automated minimization of the TE and echo spacing, absence of a phase oversampling option for the Wave-MPRAGE sequence) but were optimized to achieve similar

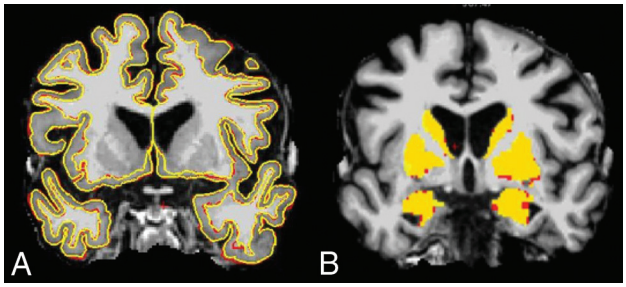

FIG 1. Subject-specific template image with superimposed pial and gray-white surface outlines (*A*) and subcortical gray matter (*B*) generated by the FreeSurfer longitudinal processing stream. The template images were generated using the standard and Wave images registered to one another. The lines representing the pial and gray-white matter surface FreeSurfer outputs on *A* and the colors on *B* are the results of subcortical gray matter segmentation. The red is the standard MPRAGE output, and yellow is the Wave-MPRAGE output. The figures demonstrate the high similarity of the segmentation generated from both sequences.

tissue contrast at the time when the center of *k*-space was acquired.

MR Imaging Protocol

For the comparative study of patients undergoing evaluation for neurodegenerative disease, MR imaging was performed on 1 of 2 clinical 3T MR scanners (Magnetom Prisma or Skyra; Siemens), using 20- or 32-channel multiarray receiver coils, depending on the fit and comfort of the patient. All clinical scans were performed using the "memory loss" protocol of our institution, which we use for the evaluation of memory loss or suspected neurodegenerative disease. In addition to the standard imaging protocol, each scan included a standard MPRAGE sequence and an ultrafast Wave-MPRAGE sequence. A summary of the MPRAGE sequence parameters is provided in Table 1.

Image Evaluation: Quantitative Comparison

The longitudinal processing stream in FreeSurfer (http://surfer.nmr.mgh.harvard.edu) was used to facilitate automated cortical segmentation and parcellation of the standard and Wave-MPRAGE images, with reduced bias for either sequence, following the approach outlined in Reuter et al.[14,15] The resulting segmentations were inspected for accuracy, but no manual edits were performed to avoid any observer bias. Regional brain volume and cortical thickness measurements were then extracted from the output of the FreeSurfer stream for the standard and Wave-MPRAGE data (Fig 1). Three cases had gross structural abnormalities or extensive motion artifacts causing failure of the FreeSurfer segmentation and were thus excluded from the quantitative analysis.

To compare the regional spatial overlap between the sequences, we coregistered the standard and Wave MPRAGE volumes using the FMRIB Linear Image Registration Tool (FLIRT; http://www.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT),[16,17] and regional spatial overlap was compared for 11 different brain regions: frontal, temporal, parietal, and occipital lobes; brain stem; cerebellum; cingulate gyrus; hippocampus; insula; cerebral white matter; and basal ganglia. We also compared the cortical thickness throughout the 33 gyri included in the FreeSurfer pipeline. Quantitative comparison was performed using 3 metrics: 1) estimation of the Pearson correlation coefficient[18] comparing the volume of all brain regions derived from the standard and Wave-MPRAGE images across the study population; 2) calculation of the absolute symmetrized percent change (ASPC) of the volume and cortical thickness measurements derived from the standard and Wave-MPRAGE images; and 3) estimation of the Dice similarity coefficient for images obtained using the 2 sequences in each brain region.

The volume (or cortical thickness) ASPC was defined as

$$ASPC = 100 * \frac{|(StandardVolume) - (WaveVolume)|}{(0.5\ [(StandardVolume) + (WaveVolume)]},$$

where "Volume" is the mean volume of the brain region of each subject.[14] The same formula was applied for the cortical thickness comparison. The ASPC is a normalized rate for the average of the volume (or cortical thickness). The ASPC does not consider the signal of the difference, the distribution of the results, or the order of processing, making it a more robust measurement than a simple percentage. The same formula was used to compare the volumes in the scan-rescan experiments, though the same sequence (standard or Wave) was used in the denominator and numerator. The results are presented with the mean and SD of the ASPC for the whole population.

The Dice coefficient of similarity compares the regional spatial overlap between the 2 sequences and is calculated as the following:[14,19,20]

$$Dice\ (Standard,\ Wave) = 2 * \frac{|Standard \cap Wave|}{(|Standard| + |Wave|)},$$

The Dice coefficient is used in the literature to estimate the regional spatial overlap, rather than just the agreement in volumetric values. It gives information on the size and structure of the analyzed region; consequently, it is a more robust way to evaluate the similarity between 2 images. The Dice coefficient ranges from 0 to 1, with a value of 1 indicating perfect spatial overlap. The Dice coefficient can be interpreted similar to the Cohen κ coefficient, as suggested by Zijdenbos et al:[21] 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and >0.80, almost perfect agreement.

To investigate voxelwise differences in gray and white matter signal intensity between the standard and Wave-MPRAGE images, we performed voxel-based morphometry (VBM) analysis of the Wave and standard MPRAGE images acquired in the 10 healthy volunteers who participated in the scan-rescan experiments. The standard and Wave-MPRAGE images were analyzed with FSLVBM[22] (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLVBM), an optimized VBM protocol[23] performed with FSL tools.[24] Brain extraction and gray matter segmentation were first performed on the structural images before registration to the Montreal Neurological Institute 152 standard space using nonlinear registration.[25] The resulting images were averaged

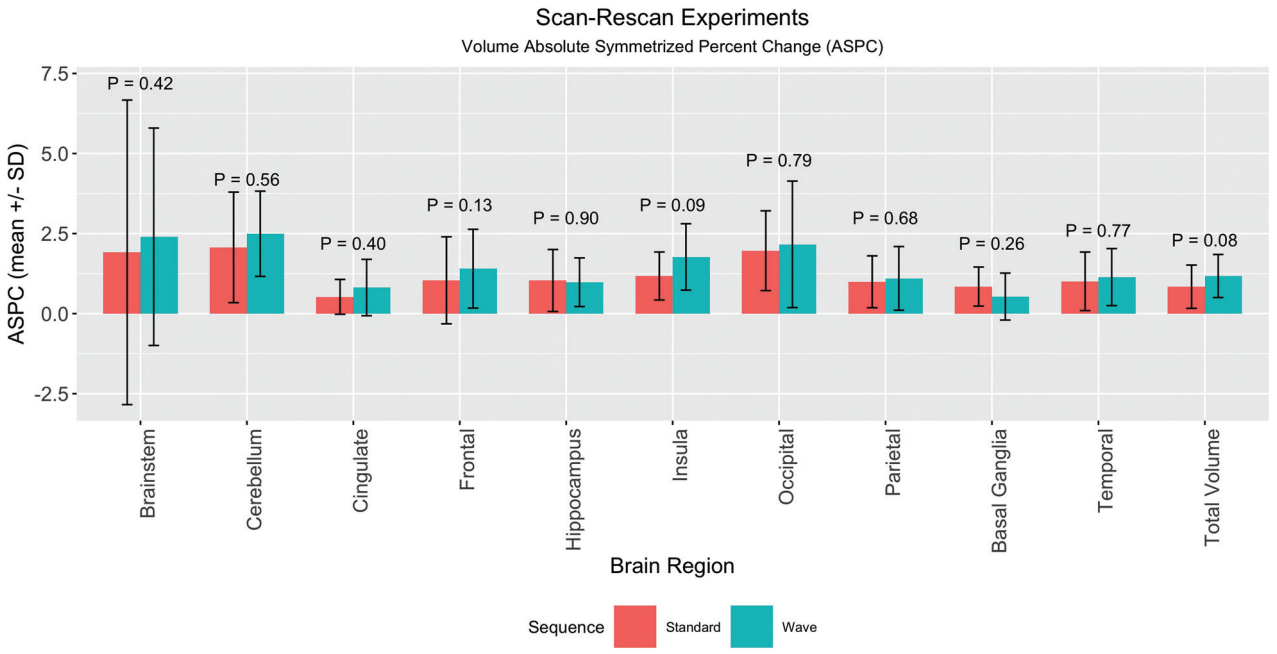

FIG 2. Barplot summarizing the results of the scan-rescan experiments in 10 healthy volunteers. The graph compares the mean and SD of the ASPC of the same sequence (ie, Wave compared with Wave, and standard compared with standard). Each sequence was acquired in different parts of the same scan session (see Materials and Methods for more details). The comparison of the Wave and Standard ASPC was not statistically different for the brain regions studied using a paired *t* test.

and flipped along the x-axis to create a left-right symmetric, study-specific gray matter template. All native gray matter images were then nonlinearly registered to this study-specific template and "modulated" to correct for local expansion/contraction due to the nonlinear spatial transformation. The modulated gray matter images were then smoothed using an isotropic Gaussian kernel with a signal of 3 mm. Finally, a voxelwise general linear model was applied using permutation-based nonparametric testing, correcting for multiple comparisons across space.

Image Evaluation: Visual Rating of Cerebral Atrophy

Following the approach described by Harper et al,[26] we performed visual evaluation of regional brain atrophy according to 6 histopathologically validated rating scales in the cohort of patients: 1) the 5-point anterior temporal scale of Davies et al[27] and Kipps et al;[28] 2) the 5-point medial temporal lobe atrophy scale of Scheltens et al;[29] 3) the 4-point posterior atrophy scale of Koedam et al;[30] and 4) the 4-point orbitofrontal, 5) anterior-cingulate, and 6) frontoinsular scales adapted by Davies et al[31] (On-line Fig 1).

Two neuroradiologists (J.C. and M.G.F.L., with 9 and 7 years of experience, respectively) were trained before performing the image analyses. To improve the interrater reliability, we selected standard anatomic landmarks for each scale, and reference images illustrating examples of each possible score on the scale were available during the evaluation, as described by Harper et al.[26] The images were anonymized for the pulse sequence information and parameters and presented in random order to the readers. Only the MPRAGE sequences from the patients' imaging examinations were evaluated.

Statistical Analysis

For the demographic and FreeSurfer data, means and SDs were calculated for the continuous variables, and percentages were calculated for the categoric variables. Paired *t* tests were used to compare the ASPCs and volumes of each cortical region. ANOVA was used to test the interaction of the head coil used (20 or 32 channels) and results of the segmentation.

For the VBM-style analyses, to test for significant differences in gray matter signal intensities between the standard and Wave-MPRAGE images, we performed voxelwise cross-subject statistical analysis using permutation-based nonparametric inference with 5000 random permutations. The results were considered significant at $P < .05$ using cluster-based threshold-free cluster enhancement,[32] corrected for multiple comparisons using the family-wise error rate.

For the visual rating of cerebral atrophy scales, we calculated the absolute difference between the scores obtained from visual evaluation of the standard MPRAGE and Wave-MPRAGE images for each brain region. The mean and maximum values of the difference between sequences was calculated. We calculated the difference in scores between the sequences (|Standard – Wave|), and between the readers (|Reader 1 – Reader 2|). We tested for equivalence of the sequences using a paired two 1-sided *t* test, with lower ($-\Delta$) and higher (Δ) bounds of −0.5 and 0.5, respectively, and a 2-sided 95% confidence interval, assuming a 2-sided significance level of 5%.[33] The bounds were based on a previous study that demonstrated clinical relevance of a mean difference greater than 0.5 comparing healthy controls with patients with dementia.[31] The two 1-sided *t* test tests whether the difference between the 2 groups is equivalent to zero and not different from two 1-sided *t* test zero. We tested whether the mean difference between the groups is statistically rejected for both the

1-sided tests, concluding that $-\Delta < |\text{Standard} - \text{Wave}| < \Delta$. In other words, we tested whether the observed effect falls within the equivalence bounds and it is small enough to consider that the groups are practically equivalent. All statistical calculations were performed using R statistical and computing software, Version 3.4.3 (http://www.r-project.org/).

RESULTS

Scan-Rescan Reliability of Standard and Wave-MPRAGE in Healthy Volunteers

Ten healthy volunteers (28.9 ± 7.1 years of age, 9 women) participated in the scan-rescan experiments. The scan-rescan reliability of brain volumes extracted using each sequence was assessed using the ASPC. The mean ASPCs derived from the Wave and standard MPRAGE images across subjects were not statistically different in any brain region (Fig 2 and On-line Table 1). The highest mean ASPCs were in the cerebellum, where the mean ASPCs for standard MPRAGE were 2.07 ± 1.73 and 2.49 ± 1.33 for Wave-MPRAGE ($P = .56$). The lowest ASPCs were in the cingulate cortex, where the mean ASPCs for standard MPRAGE were 0.52 ± 0.54 and 0.81 ± 0.88 for Wave-MPRAGE ($P = .40$). The highest mean difference between the ASPCs was in the cerebellum, 1.83 ± 1.16, and the lowest was in the frontal lobes, 0.63 ± 0.40. We did not find an interaction of the segmentation results with the use of either the 20- or 32- channel coils ($P > .05$).

Comparative Evaluation of Standard and Wave-MPRAGE in Neurodegenerative Disease

Forty-three consecutive adults participated in the clinical comparative evaluation of the standard and Wave-MPRAGE sequences. Demographic information, including age, sex, and clinical indication for undergoing MR imaging, are described in Table 2. Of the 43 subjects, 3 (7.0%) were excluded due to the presence of structural abnormalities or severe motion degradation resulting in failure of the FreeSurfer segmentation. Quantitative evaluation of brain tissue volumes showed excellent agreement between the 2 sequences, with Dice coefficients corresponding to almost perfect agreement for all of the brain regions evaluated (κ ranged from 0.86 to 0.95) (On-line Tables 2 and 3 and Fig 3). The regions with the lowest mean Dice coefficients were the

Table 2: Clinical characteristics of the patients

Characteristics	Total ($n = 43$)	Included in Quantitative Analysis ($n = 40$)[a]
Male (%)	27 (62.8%)	25 (62.5%)
Age (median) (range) (yr)	72 (18–86)	70 (18–86)
20-Channel coil (%)	19 (46.3%)	17 (37.8%)
Research study (%)	12 (27.9%)	11 (24.4%)
Study indication		
Cognitive impairment	18 (41.9%)	18 (45.0%)
Chronic white matter disease	8 (18.6%)	7 (17.5%)
Parkinsonism	5 (11.6%)	5 (12.5%)
Stroke	5 (11.6%)	4 (10.0%)
Traumatic brain injury	3 (7.0%)	2 (5.0%)
Ataxia	2 (4.7%)	2 (5.0%)
Epilepsy	1 (2.3%)	1 (2.5%)
Mitochondrial disease	1 (2.3%)	1 (2.5%)

[a] Three subjects were excluded from the FreeSurfer quantitative analysis due to large structural changes and/or severe motion artifact resulting in failure of the automated FreeSurfer longitudinal processing stream.

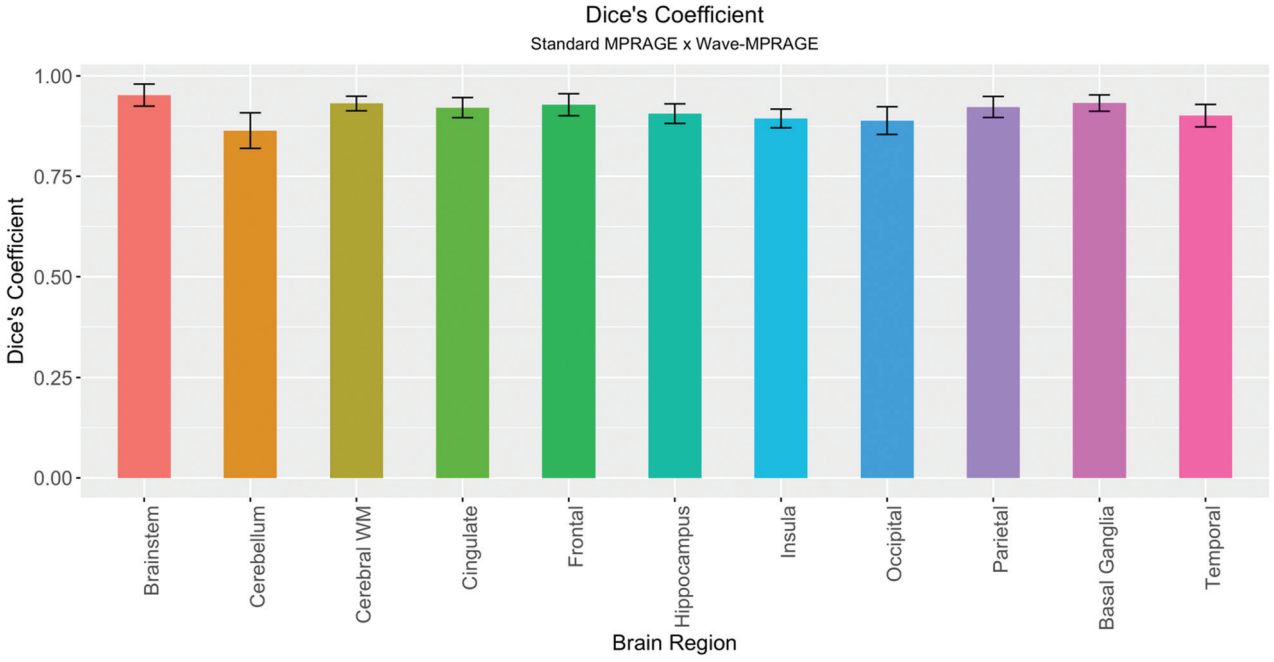

FIG 3. Barplot demonstrating the mean and SD of the Dice coefficients over the brain regions evaluated, testing the spatial overlap between the standard and Wave MPRAGE sequences. All the coefficients demonstrated almost perfect ($\kappa > 0.85$) overlap.

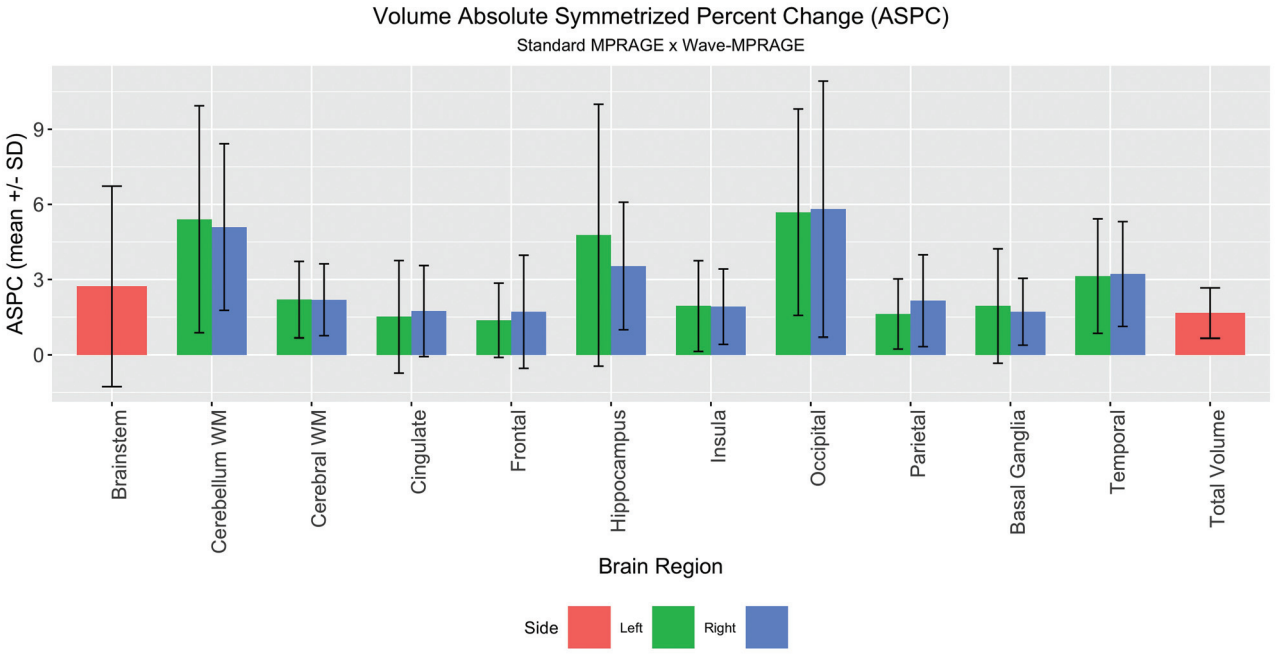

FIG 4. Barplot demonstrating the mean and SD of the ASPC for the volumes of the studied brain regions. The largest differences were in the occipital lobes and cerebellar white matter.

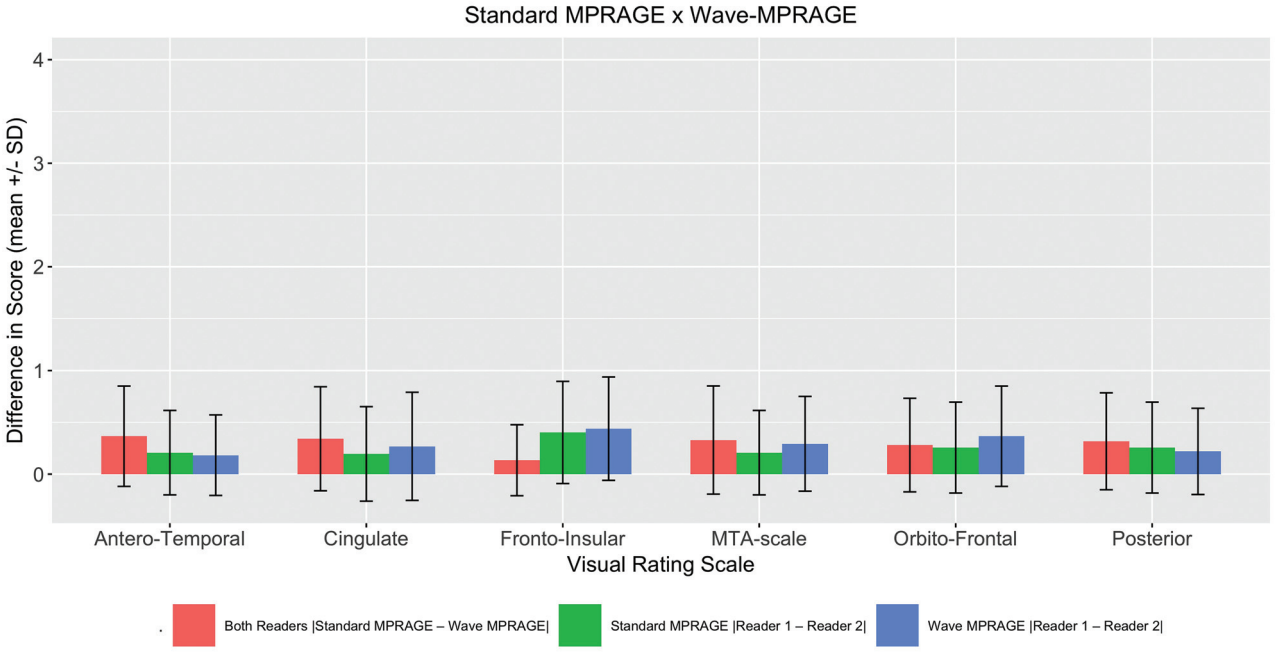

FIG 5. Barplot demonstrating the absolute values of the differences in the visual rating scores (mean ± SD) for the Wave MPRAGE and standard MPRAGE sequences (*red bars*). Differences between the 2 raters for the standard MPRAGE images (*green bars*) and Wave MPRAGE images (*blue bars*) are also provided. MTA indicates medial temporal lobe.

cerebellum (0.86 ± 0.04), insula (0.89 ± 0.02), and occipital lobes (0.89 ± 0.03).

The Pearson correlation coefficient demonstrated a high degree of correlation when comparing brain volumes between the standard and Wave-MPRAGE sequences, ranging from 0.90 to 0.99 ($P < .001$) (On-line Tables 2 and 3 and On-line Fig 2). The mean ASPC ranged from 1.4% ± 1.5% to 5.8% ± 5.1% (On-line Table 2). The areas with the highest ASPC were the same areas with the lowest Dice coefficients: left and right cerebellar hemispheres (with mean ASPCs of 5.4% ± 4.5% and 5.1% ± 3.3%, respectively; and a mean Dice coefficient of 0.86 ± 0.04) and left and right occipital lobes (with mean ASPCs of 5.7% ± 4.1% and 5.8% ± 5.1%, respectively; and a mean Dice coefficient of 0.89 ± 0.03) (On-line Table 2 and Fig 4). The area with highest mean Dice coefficient was the brain stem (0.95 ± 0.03) with a mean ASPC of 2.7% ± 4.0%. Similar results were observed for

the mean ASPC for cortical thickness, ranging from 1.4% ± 1.1% in the left superior parietal gyrus to 4.8% ± 4.4% in the right pericalcarine gyrus (On-line Figs 3 and 4).The average of total brain volume was 1095.7 ±123.4 mL for the standard MPRAGE and 1077.9 ±119.9 mL for the Wave-MPRAGE, which corresponds to a mean ASPC of 1.66 ± 1.0% (Fig 4).

A VBM-style comparison of the standard and Wave-MPRAGE images in the healthy volunteers participating in the scan-rescan experiments revealed significant differences in signal intensities in the bilateral occipital lobes, cingulate gyri, cerebellar hemispheres, and hippocampi (On-line Fig 5). These regions corresponded to the areas showing the most significant differences in the ASPC in the brain volume quantitative evaluation (On-line Table 2).

In the visual rating of cerebral atrophy, the maximum difference between the atrophy scores for the standard MPRAGE and Wave-MPRAGE images on any of the rating scales was 1 point. The average differences between the scores for the standard MPRAGE and Wave-MPRAGE images are provided in Fig 5. The highest mean difference was in the frontoinsular scale (0.40 ± 0.5 and 0.44 ± 0.5 for the standard and Wave MPRAGE sequences, respectively). The mean difference in scores |Standard – Wave| was equivalent to zero for all the scales (two 1-sided t test $P < .05$) (On-line Fig 4).

DISCUSSION

This study compared a highly accelerated Wave-MPRAGE sequence with the standard MPRAGE acquisition for brain volume assessment. We included an assessment of scan-rescan reliability of the standard and Wave-MPRAGE sequences in healthy volunteers and a subsequent comparative evaluation of the standard and Wave-MPRAGE sequences in a cohort of patients undergoing investigation of neurodegenerative disease, which included consecutive MR imaging examinations from both clinical and research settings. We used different quantitative metrics to compare the sequences: the Pearson correlation coefficient, ASPC, and Dice similarity coefficient. Each measure provides different information. The Pearson correlation coefficient demonstrates the data variation. However, it does not consider the magnitude of the difference between the 2 datasets, which was evaluated using the ASPC. The Dice coefficient compares the within- and between-subject variability, as well as the regional spatial overlap for different brain regions. In addition to these computed metrics, we also included evaluation of the images using a visual rating scale to determine whether any difference in image quality using Wave-MPRAGE would impact a radiologist's visual evaluation of brain volume, which is often an important part of the patient's clinical evaluation.

On visual inspection, Wave-MPRAGE images showed greater image noise than standard MPRAGE, particularly in the central brain (On-line Fig 1). This is an expected finding as there is less SNR in the center of the coil in comparison to the periphery, and the SNR scales with the square root of the acceleration factor, R.[12] Several strategies were implemented in a previous technical development study[34] to minimize both noise amplification and Wave-specific blurring artifacts. In this prior study, it was noted that by means of the optimized Wave-MPRAGE sequence, it is possible to decrease the g-factor by up to 1.10 times with an acceleration factor of $R = 3 \times 3$. Moreover, a subjective evaluation by multiple raters comparing both sequences, standard and Wave-MPRAGE, demonstrated comparable image quality between them.[34] In our study, all of the metrics evaluated also showed high similarity between the 2 sequences, suggesting that a mild increase in image noise with Wave-MPRAGE had a negligible impact on quantitative brain tissue segmentation and qualitative brain volume evaluation. Nevertheless, we acknowledge that the greater noise in the Wave-MPRAGE images may affect the clinical assessment made by radiologists, which is an important consideration that should be addressed and further refined in the translation of Wave-MPRAGE to the clinic. Effort is currently underway to incorporate postprocessing techniques such as denoising and image regularization into the pipeline, which can be tuned to improve the noise in the Wave-MPRAGE images without incurring excessive spatial blurring.

In our scan-rescan experiments, we did not find a statistically significant difference in mean ASPCs between the 2 sequences for any brain region. Furthermore, the mean ASPCs for each sequence in the scan-rescan experiments were consistently smaller than the corresponding mean ASPCs between Wave and standard MPRAGE in the patients undergoing evaluation for neurodegenerative disorders, attesting to the relatively low intra-individual variability of brain volumes estimated from the Wave and standard MPRAGE acquisitions. We subsequently evaluated Wave-MPRAGE in a clinical population including patients undergoing evaluation for suspected neurodegenerative disease, which included a range of both normal and abnormal brain tissue volumes more representative of what would be encountered in clinical practice. The correlation coefficients and ASPC for the brain volume and cortical thickness between the standard and Wave-MPRAGE sequences in our study were comparable with those obtained in other studies that performed test and retest analysis in more homogeneous cohorts composed of healthy volunteers.[14,18,35,36] In addition, we demonstrated almost perfect correlation in our data for the voxel-by-voxel spatial overlap comparison for all brain regions.[14]

The test and retest comparison of the FreeSurfer longitudinal stream was evaluated by Reuter et al[14] when the algorithm was launched, comparing the brain volumes of healthy volunteers acquired in the same scan session, using the same sequence. The cortical thickness and volume ASPC results in the article by Reuter et al were approximately 2%–3% for most brain regions, similar to our cortical thickness ASPC results. In the brain volume analysis, we found that most of the regions also had an ASPC of approximately 2%–3%, except for the occipital lobes (approximately 5%–6%). Most interesting, the cuneus was the region with the worst performance in the test and retest analysis using the cross and longitudinal streams in the original article as well.[14] Given that the visual cortex is one of the thinnest cortical regions with a high degree of myelination, the lower gray-white matter contrast in this region could explain the poorer performance of the FreeSurfer segmentation in this area. Alternatively, small differences between the Wave-MPRAGE and standard MPRAGE acquisitions could contribute to the small differences in brain tissue volumes observed in this study (with ASPC

ranging from 1.4% to 5.8%, depending on brain region). These differences are similar to those observed in other studies evaluating the effect of small changes in the MPRAGE parameters,[37] scanner model,[38] FreeSurfer version, and operating system,[39] underscoring the importance of standardizing the acquisition and data processing procedures for longitudinal studies.

In our scan-rescan experiments, the mean ASPCs for brain volumes derived from the standard and Wave-MPRAGE images were not statistically different. The comparisons performed in the scan-rescan experiments were different from the comparisons in the cohort of patients. In this experiment, we compared the Wave sequence with the Wave sequence and the standard sequence with the standard sequence, demonstrating high interscan reliability. Moreover, the scan-rescan experiments were performed in healthy volunteers in a more controlled environment, resulting in fewer artifacts including fewer motion artifacts. The scan-rescan experiment reinforces the notion that when fewer variables are present, the differences will tend to be closer to zero.

In the Dice coefficient of similarity comparison of the regional overlap for different brain regions, the cerebellum had the lowest score (0.86), though it is still interpreted as almost perfect agreement. We believe that the main reason for the differences found in the spatial overlap is that FreeSurfer fails to remove the cerebellar peduncle and venous sinuses accurately. In On-line Fig 6, there is an example from one of our cases showing the FreeSurfer segmentation. The software extends the segmentation to the confluence of the transverse sinuses in this case, leading to inaccurate segmentation. Additionally, the signal intensities of the cerebellar white matter and the peduncles are very similar, making the segmentation more difficult. These limitations lead to more errors and subsequent differences in the Dice coefficient, as noted in our Results.[40]

In the visual rating analysis, the mean score differences between the standard and Wave-MPRAGE sequences were small. The scales used here have a well-known intrinsic and subjective variability, with interrater agreement ranging from moderate agreement to almost perfect agreement in a larger cohort study.[26] Therefore, the score differences are more likely related to the intrinsic variability of the scales than to the sequence used. Moreover, in the study of Harper et al,[26] the differences between the mean scores of healthy controls compared with patients with dementia were mostly greater than 0.5. Consequently, differences inferior to this value (as observed in our study) are unlikely to be clinically significant. These results suggest that Wave-MPRAGE could potentially replace standard MPRAGE in clinical brain imaging protocols, resulting in more efficient use of MR imaging resources, noting that more validation studies are needed to demonstrate the utility and comparability of Wave-MPRAGE with standard MPRAGE in aiding the clinical diagnosis of neurodegenerative disorders.

This study has several limitations. To prevent selection bias, we included consecutive brain volumes acquired in the study period for volume-loss evaluation. However, a few cases ($n = 3$) had gross structural abnormalities resulting in failure of the FreeSurfer segmentation and therefore had to be excluded from the quantitative volumetric analysis. We have a relatively small sample size, though our number is still larger than that in many studies making similar test and retest comparisons.[14,36,37] Most of these studies included healthy subjects, while our cohort included patients undergoing clinical evaluation for neurodegenerative disease, increasing the complexity of acquiring the images (eg, due to greater patient motion). Moreover, we also used strategies to decrease the error (bias) in the volume evaluations, specifically the FreeSurfer longitudinal stream. Although the radiologist observers were blinded to the pulse sequence used, some features in the images could help in the identification of the sequence. The image quality of the Wave-MPRAGE and other Wave sequences has been evaluated in a recently published work[34] and was not formally evaluated here, because the goal of our study was to evaluate the impact of sequence selection on quantitative and semiquantitative brain volume estimation.

CONCLUSIONS

Wave-MPRAGE provided reliability similar to that of standard MPRAGE for regional evaluation of brain atrophy using automated segmentation of brain tissue volumes, cortical thickness measurements, and semiquantitative visual rating scales, despite a 3- to 4-fold decrease in acquisition time. Adoption of Wave-MPRAGE for volumetric imaging of patients with suspected neurodegenerative disease could provide more efficient use of MR imaging resources in both clinical and research settings.

Disclosures: Maria Gabriela F. Longo—*RELATED*: *Consulting Fee or Honorarium*: Siemens. Stephen F. Cauley—*RELATED*: *Grant*: Siemens, *Comments*: I receive research support from Siemens; *UNRELATED*: *Grants/Grants Pending*: Siemens, *Comments*: I receive research support from Siemens beyond the scope of this work. Kawin Setsompop—*RELATED*: *Grant*: National Institutes of Health, *Comments*: R01EB020613.* Daniel Polak—*OTHER RELATIONSHIPS*: My PhD salary is paid by Siemens. Daniel Splitthoff—*RELATED*: *Other*: The work was supported from side while being an employee of Siemens; *UNRELATED*: *Employment*: The work was supported from site while being an employee of Siemens. Wei Liu—*UNRELATED*: *Employment*: Siemens Shenzhen Magnetic Resonance Ltd. Otto Rapalino—*UNRELATED*: *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: GE Healthcare, *Comments*: travel expenses for medical conference in San Diego 2019, role: speaker. Susie Y. Huang—*RELATED*: *Grant*: Siemens, *Comments*: research grant.* Maya Polackal—*UNRELATED*: *Employment*: Athinoula A. Martinos Center, *Comments*: research assistant. *Money paid to the institution.

REFERENCES

1. Rusinek H, de Leon MJ, George AE, et al. **Alzheimer disease: measuring loss of cerebral gray matter with MR imaging.** *Radiology* 1991;178:109–14 CrossRef Medline
2. Lehmann M, Douiri A, Kim LG, et al. **Atrophy patterns in Alzheimer's disease and semantic dementia: a comparison of FreeSurfer and manual volumetric measurements.** *Neuroimage* 2010;49:2264–74 CrossRef Medline
3. Frisoni GB, Fox NC, Jack CR, et al. **The clinical use of structural MRI in Alzheimer disease.** *Nat Rev Neurol* 2010;6:67–77 CrossRef Medline
4. Zarei M, Ibarretxe-Bilbao N, Compta Y, et al. **Cortical thinning is associated with disease stages and dementia in Parkinson's disease.** *J Neurol Neurosurg Psychiatry* 2013;84:875–82 CrossRef Medline
5. Mugler JP, Brookeman JR. **Rapid three-dimensional T1-weighted MR imaging with the MP-RAGE sequence.** *J Magn Reson Imaging* 1991;1:561–67 CrossRef Medline
6. Munn Z, Pearson A, Jordan Z, et al. **Patient anxiety and satisfaction in a magnetic resonance imaging department: Initial results from an action research study.** *J Med Imaging Radiat Sci* 2015;46:23–29 CrossRef Medline

7. Havsteen I, Ohlhues A, Madsen KH, et al. **Are movement artifacts in magnetic resonance imaging a real problem? A narrative review.** *Front Neurol* 2017;8:232 CrossRef Medline
8. Savalia NK, Agres PF, Chan MY, et al. **Motion-related artifacts in structural brain images revealed with independent estimates of in-scanner head motion.** *Hum Brain Mapp* 2017;38:472–92 CrossRef Medline
9. Bilgic B, Gagoski BA, Cauley SF, et al. **Wave-CAIPI for highly accelerated 3D imaging.** *Magn Reson Med* 2015;73:2152–62 CrossRef Medline
10. Cauley SF, Setsompop K, Bilgic B, et al. **Autocalibrated wave-CAIPI reconstruction; Joint optimization of k-space trajectory and parallel imaging reconstruction.** *Magn Reson Med* 2017;78:1093–99 CrossRef Medline
11. Conklin J, Longo MGF, Cauley SF, et al. **Validation of highly accelerated Wave-CAIPI SWI compared with conventional SWI and T2*-weighted gradient recalled-echo for routine clinical brain MRI at 3T.** *AJNR Am J Neuroradiol* 2019;40:2073–80 CrossRef Medline
12. Polak D, Setsompop K, Cauley SF, et al. **Wave-CAIPI for highly accelerated MP-RAGE imaging.** *Magn Reson Med* 2018;79:401–06 CrossRef Medline
13. Nielsen JA, Mair RW, Baker JT, et al. **Precision brain morphometry: feasibility and opportunities of extreme rapid scans.** *bioRxiv* January 26, 2019 https://www.biorxiv.org/content/10.1101/530436v1. Accessed May 1, 2020
14. Reuter M, Schmansky NJ, Rosas HD, et al. **Within-subject template estimation for unbiased robust and sensitive longitudinal image analysis.** *Neuroimage* 2012;61:1402–18 CrossRef Medline
15. Reuter M, Rosas HD, Fischl B. **Highly accurate inverse consistent robust registration: a robust approach.** *Neuroimage* 2010;53:1181–96 CrossRef Medline
16. Jenkinson M, Bannister P, Brady M, et al. **Improved optimization for the robust and accurate linear registration and motion correction of brain images.** *Neuroimage* 2002;17:825–41 CrossRef Medline
17. Jenkinson M, Smith S. **A global optimisation method for robust affine registration of brain images.** *Med Image Anal* 2001;5:143–56 CrossRef Medline
18. Steen RG, Hamer RM, Lieberman JA. **Measuring brain volume by MR imaging: Impact of measurement precision and natural variation on sample size requirements.** *AJNR Am J Neuroradiol* 2007;28:1119–25 CrossRef Medline
19. Ranta ME, Chen M, Crocetti D, et al. **Automated MRI parcellation of the frontal lobe.** *Hum Brain Mapp* 2014;35:2009–26 CrossRef Medline
20. Dice L. **Measurement of the amount of ecological association between species.** *Ecology* 1945;26:297–302 CrossRef
21. Zijdenbos AP, Dawant BM, Margolin RA, et al. **Morphometric analysis of white matter lesions in MR images: method and validation.** *IEEE Trans Med Imaging* 1994;13:716–24 CrossRef Medline
22. Douaud G, Smith S, Jenkinson M, et al. **Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia.** *Brain* 2007;130:2375–86 CrossRef Medline
23. Good CD, Johnsrude IS, Ashburner J, et al. **A voxel-based morphometric study of ageing in 465 normal adult human brains.** *Neuroimage* 2001;14:21–36 CrossRef Medline
24. Smith SM, Jenkinson M, Woolrich MW, et al. **Advances in functional and structural MR image analysis and implementation as FSL.** *Neuroimage* 2004;23(Suppl 1):1S208–19 CrossRef Medline
25. Andersson JL, Jenkinson M. Smith S. **Non-linear registration, aka spatial normalization**; **FMRIB technical report TR07JA2.** 2007 https://www.fmrib.ox.ac.uk/datasets/techrep/tr07ja2/tr07ja2.pdf. Accessed May 1, 2020
26. Harper L, Fumagalli GG, Barkhof F, et al. **MRI visual rating scales in the diagnosis of dementia: evaluation in 184 post-mortem confirmed cases.** *Brain* 2016;139:1211–25 CrossRef Medline
27. Davies RR, Kipps CM, Mitchell J, et al. **Progression in frontotemporal dementia: identifying a benign behavioral variant by magnetic resonance imaging.** *Arch Neurol* 2006;63:1627–31 CrossRef Medline
28. Kipps CM, Davies RR, Mitchell J, et al. **Clinical significance of lobar atrophy in frontotemporal dementia: application of an MRI visual rating scale.** *Dement Geriatr Cogn Disord* 2007;23:334–42 CrossRef Medline
29. Scheltens P, Leys D, Barkhof F, et al. **Atrophy of medial temporal lobes on MRI in "probable Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates.** *J Neurol Neurosurg Psychiatry* 1992;55:967–72 CrossRef Medline
30. Koedam E, Lehmann M, van der Flier WM, et al. **Visual assessment of posterior atrophy development of a MRI rating scale.** *Eur Radiol* 2011;21:2618–25 CrossRef Medline
31. Davies RR, Scahill VL, Graham A, et al. **Development of an MRI rating scale for multiple brain regions: comparison with volumetrics and with voxel-based morphometry.** *Neuroradiology* 2009;51:491–503 CrossRef Medline
32. Smith SM, Nichols TE. **Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference.** *Neuroimage* 2009;44:83–98 CrossRef Medline
33. Ahn S, Park SH, Lee KH. **How to demonstrate similarity using noninferiority and equivalence statistical testing in radiology research.** *Radiology* 2013;267:328–38 CrossRef Medline
34. Polak D, Cauley S, Huang SY, et al. **Highly-accelerated volumetric brain examination using optimized wave-CAIPI encoding.** *J Magn Reson Imaging* 2019;50:961–14 CrossRef Medline
35. Maclaren J, Han Z, Vos SB, et al. **Reliability of brain volume measurements: a test-retest dataset.** *Sci Data* 2014;1:1–9 CrossRef Medline
36. Iscan Z, Jin TB, Kendrick A, et al. **Test-retest reliability of FreeSurfer measurements within and between sites: effects of visual approval process.** *Hum Brain Mapp* 2015;36:3472–85 CrossRef Medline
37. Haller S, Falkovskiy P, Meuli R, et al. **Basic MR sequence parameters systematically bias automated brain volume estimation.** *Neuroradiology* 2016;58:1153–60 CrossRef Medline
38. 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;46:177–92 CrossRef Medline
39. Gronenschild E, Habets P, Jacobs HIL, et al. **The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements.** *PLoS One* 2012;7:e38234 CrossRef Medline
40. Hwang J, Kim J, Han Y, et al. **An automatic cerebellum extraction method in T1-weighted brain MR images using an active contour model with a shape prior.** *Magn Reson Imaging* 2011;29:1014–22 CrossRef Medline