HHS Public Access

Author manuscript

Neuroimage Rep. Author manuscript; available in PMC 2022 August 25.

Published in final edited form as:

Neuroimage Rep. 2022 June ; 2(2): . doi:10.1016/j.ynirp.2022.100086.

Test-retest reliability of FreeSurfer-derived volume, area and cortical thickness from MPRAGE and MP2RAGE brain MRI images

Graham N. Knussmann[a], **Jeffrey S. Anderson**[a], **Molly B.D. Prigge**[a], **Douglas C. Dean III**[b,c,d], **Nicholas Lange**[e], **Erin D. Bigler**[f,g,h,i], **Andrew L. Alexander**[b,d,j], **Janet E. Lainhart**[b,j], **Brandon A. Zielinski**[a,g,k], **Jace B. King**[a,*]

[a]Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, UT, USA

[b]Waisman Center, University of Wisconsin-Madison, Madison, WI, USA

[c]Department of Pediatrics, University of Wisconsin-Madison, Madison, WI, USA

[d]Department of Medical Physics, University of Wisconsin-Madison, Madison, WI, USA

[e]Department of Psychiatry, Harvard Medical School, Boston, MA, USA

[f]Department of Psychology and Neuroscience Center, Brigham Young University, Provo, UT, USA

[g]Department of Neurology, University of Utah, Salt Lake City, UT, USA

[h]Department of Psychiatry, University of Utah, Salt Lake City, UT, USA

[i]Department of Neurology, University of California-Davis, Davis, CA, USA

[j]Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, USA

[k]Department of Pediatrics, University of Utah, Salt Lake City, UT, USA

Abstract

[*]Corresponding author. University of Utah, 729 Arapeen Drive, Salt Lake City, UT, 84108, USA. jace.king@hsc.utah.edu (J.B. King).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT author statement

Graham N. Knussmann: Formal analysis, Investigation, Software, Validation, Visualization, Writing – Original Draft. Jeffrey S. Anderson: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing. Molly B. D. Prigge: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Supervision, Writing – Original Draft, Writing – Review & Editing. Douglas C. Dean III: Investigation, Methodology, Writing – Review & Editing. Nicholas Lange: Conceptualization, Funding Acquisition, Methodology, Writing – Review & Editing. Erin D. Bigler: Conceptualization, Funding Acquisition, Supervision, Writing – Review & Editing. Andrew L. Alexander: Conceptualization, Funding Acquisition, Methodology, Project Administration, Supervision, Writing – Review & Editing. Janet E. Lainhart: Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Review & Editing. Brandon A. Zielinski: Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Review & Editing. Jace B. King: Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Software, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing.

Appendix A.: Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ynirp.2022.100086.

Background and purpose: Large MRI studies often pool data gathered from widely varying imaging sequences. Pooled data creates a potential source of variation in structural analyses which may cause misinterpretation of findings. The purpose of this study is to determine if data acquired using different scan sequences, head coils and scanners offers consistent structural measurements.

Materials and methods: Participants (163 right-handed males: 82 typically developing controls, 81 participants with autism spectrum disorder) were scanned on the same day using an MPRAGE sequence with a 12-channel headcoil on a Siemens 3T Trio scanner and an MP2RAGE sequence with a 64-channel headcoil on a Siemens 3T Prisma scanner. Segmentation was performed using FreeSurfer to identify regions exhibiting variation between sequences on measures of volume, surface area, and cortical thickness. Intraclass correlation coefficient (ICC) and mean percent difference (MPD) were used as test-retest reproducibility measures.

Results: ICC for total brain segmented volume yielded a 0.99 intraclass correlation, demonstrating high overall volumetric reproducibility. Comparison of individual regions of interest resulted in greater variation. Volumetric variability, although low overall, was greatest in the entorhinal cortex (ICC = 0.71), frontal (ICC = 0.60) and temporal (ICC = 0.60) poles. Surface area variability was greatest in the insula (ICC = 0.65), temporal (ICC = 0.64) and frontal (ICC = 0.68) poles. Cortical thickness was most variable in the frontal (ICC = 0.41) and temporal (ICC = 0.35) poles.

Conclusion: Data collected on different scanners and head coils using MPRAGE and MP2RAGE are generally consistent for surface area and volume estimates. However, regional variability may constrain accuracy in some regions and cortical thickness measurements exhibit higher generalized variability.

Keywords

MPRAGE; MP2RAGE; Reliability; FreeSurfer

1. Introduction

Volumetric gray matter quantitation of the brain has been widely used in studies of neurodegenerative, (Frisoni, Fox, Jack, Scheltens, Thompson) neurological (Eshaghi et al., 2018), psychiatric (Koshiyama et al., 2018), and neurodevelopmental (Lange, Travers, Bigler, et al.; Zielinski, Prigge, Nielsen, et al.) disorders with the goal of understanding the neurobiology of the disorders and mechanisms involved. Multisite datasets have emerged as powerful tools for discovery and hypothesis-based analyses of changes in brain development (Di Martino, 2012; Jack et al., 2008; van Erp et al., 2018), but data are often acquired on different scanners (between and within scanner manufacturers), at different times, or at different institutions with heterogenous acquisition parameters and techniques. One of the largest obstacles in magnetic resonance imaging (MRI) research today is the number of technical variables that can impact the outcome of study results. For example, effects of software version, workstation type and operating system have been shown to impact the reproducibility of analysis (Gronenschild et al., 2012). The successful use of pooled data for MRI studies involving human research depends on standardized data structures and quality control measures that facilitate reproducibility across datasets.

MPRAGE, or 3D-MPRAGE (three-dimensional magnetization-prepared rapid gradient-echo), is an imaging sequence commonly used to capture T1-weighted 3D data with good signal to noise ratio and high resolution in a short imaging time (Mugler, Brookeman). An extension of MPRAGE was later developed, MP2RAGE, which combines two gradient echo images with different inversion times to provide higher quality T1-weighted images by removing signal noise due to proton density contrast, B1 inhomogeneity effects, T2* contrast, and reception field bias (Marques, Kober, Krueger, van der Zwaag, Van de Moortele, Gruetter). While T1-weighted imaging using MPRAGE has long been a mainstay of medical imaging research, recent studies have increasingly taken advantage of the improved gray-white contrast resolution of MP2RAGE sequences (Kecskemeti, Samsonov, Hurley, Dean, Field, Alexander; Okubo, Okada, Yamamoto, et al.). Thus, researchers using pooled or longitudinal datasets are now confronted with mixed datasets containing both MPRAGE and MP2RAGE images.

Automated image analysis software has facilitated streamlined approaches for rapidly analyzing many participants to identify subtle alterations in brain volume, surface area and cortical thickness. Researchers are now investigating how the structural output resulting from images acquired using MPRAGE versus MP2RAGE differs using various automated processing streams. Improved deep gray matter contrast and more reliable volumes were found in MP2RAGE versus MPRAGE images processed with SPM12 in a group of adult controls (n = 20) (Okubo, Okada, Yamamoto, et al.). More recently, Alonso and colleagues compared MPRAGE vs MP2RAGE volumes (gray matter, white matter, CSF) derived from three processing pipelines (SPM, FSL's SIENA/X, MorphoBox) in a group of control participants (n = 24) (Alonso et al., 2020). They found increased volumes from the MP2RAGE images and cautioned combining MPRAGE and MP2RAGE images in the same analysis.

Another widely used processing pipeline that provides estimates of cortical thickness, surface area and cortical and subcortical volumes is FreeSurfer. FreeSurfer's morphometric procedures have demonstrated good test-retest reliability across scanner manufacturers and across field strengths (Han et al., 2006; Reuter et al., 2012), but few studies have investigated repeatability associated with MP2RAGE compared to other MPRAGE images. A small FreeSurfer study of healthy participants (n = 8) showed thinner cortical surface estimates resulting from MP2RAGE versus Multi-echo MPRAGE (MEMPRAGE) images (Fujimoto et al., 2014). Recently, a test-retest variability study investigated the effect of scanner headcoil in MPRAGE, MEMPRAGE, and MP2RAGE images acquired during the same scan session in 24 participants, and reliable subcortical estimates from MP2RAGE were found, yet a comparison of brain volumes between the sequences was not performed (Yan, Qian, Marechal, et al., 2020).

The purpose of this study is to expand on previous studies and determine the within-subject reproducibility of FreeSurfer derived volume, cortical thickness, and surface area measures from two scan protocols, MPRAGE and MP2RAGE, acquired on two different Siemens 3T MR scanners (Trio and Prisma) and with two different head coils in a sample of typically developing participants (TDC; n = 82) and a clinical sample with autism spectrum disorder

(ASD; n = 81). Thus, our findings are relevant to real-world multisite datasets, with different scanners and protocols, and examine reproducibility in a clinical population.

2. Material and methods

2.1. Study demographics

The data for this study were acquired as part of an ongoing longitudinal study of ASD compared to TDC. All study procedures were approved by the Institutional Review Board and participants provided written informed assent or consent prior to study participation with parental or legal guardian consent required of all subjects younger than 18 years. The subjects included in this analysis are a subset of the total sample of ASD and TDC male youth and adults in the longitudinal study who underwent scanning on both the Siemens Trio and Prisma scanners and were right-handed. The 81 ASD and 82 TDC participants were well-matched for age (ASD mean age = 27.3 years, range = 14.7–57.9; TDC mean age 27.3 years, range: 16.3–46.9; t = 0.04, p = .9).

2.2. MRI scanning protocol

All study participants were scanned using the MPRAGE sequence on a Siemens 3T Trio scanner (TR = 2300 ms, TE = 2.91 ms, TI = 900 ms, flip angle = 9°, field of view = 256 mm, 160 slices, resolution = 1.2 × 1 × 1 mm, 12-channel headcoil, sagittal acquisition) and scanned again on the same day using the MP2RAGE sequence on a Siemens 3T Prisma scanner (TR = 5000 ms, TE = 2.93 ms, TI_1 = 700 ms, TI_2 = 2030 ms, flip $angle_1$ = 4°, flip $angle_2$ = 5°, field of view = 256 mm, 176 slices, resolution = 1 × 1 × 1 mm, 64-channel headcoil, sagittal acquisition).

2.3. Data processing

Segmentation was performed on one workstation using the FreeSurfer pipeline (Version 6.0.0; Linux-centos6_x86_64) on images from each scan (Dale, Fischl, Sereno; Fischl, Dale; Fischl, Salat, Busa, et al.; Fischl, Sereno, Dale; Fischl, van der Kouwe, Destrieux, et al.). Background noise in the MP2RAGE was removed prior to completing the FreeSurfer pipeline. This was completed by multiplying the uniform image by the second inverse image. Measurements of interest included cortical surface area, mean cortical thickness, and cortical volume parcellated into 34 bilateral regions of interest (ROIs) (Desikan et al., 2006), and total brain volume. An example of an MPRAGE, an MP2RAGE, and the resulting FreeSurfer output for a single subject can be found in the Supplementary Material (Supplementary Material Fig. 1).

All processed images were visually inspected by a trained technician for scan artifacts, movement distortion, and structural anomalies. The initial dataset contained 179 participants, one of which was excluded due to incomplete dataset. Eleven scans from 10 participants were excluded from the study due to artifacts, movement error or structural anomalies. Five female participants were then removed resulting in a final sample size of 163 male subjects. To represent imaging studies using large datasets where manual editing is not feasible, our findings summarize raw FreeSurfer output without any manual intervention or reprocessing.

2.4. Statistical analysis

Statistical analyses were performed in the MATLAB computing environment (MathWorks, Natick, MA, USA). Intraclass correlation coefficient (ICC(A,1)) (McGraw and Wong, 1996; Salarian, 2019) was calculated as a measure of reproducibility (poor: <0.50, moderate: 0.50-0.75, good: 0.75-0.90, excellent >0.90) for FreeSurfer derived area, volume and thickness for each ROI in the Desikan-Killiany atlas (n = 34 regions per hemisphere). Mean percent difference (MPD) was also used to compare each ROI from data acquired with MPRAGE and MP2RAGE. This calculation was performed for volume, surface area and cortical thickness on each ROI as well as global brain measures for every individual in the sample. The resulting percent difference within each subject's scans was then aggregated to arrive at a mean percent difference per ROI.

$$\sum_{n=1}^{163}\left(\frac{|MPRAGE - MP2RAGE|}{\frac{MPRAGE + MP2RAGE}{2}}\right) \times 100$$

ICC and MPD were calculated for each ROI separately in the left and right hemisphere. Primary results are presented on bilateral (averaged ICC or MPD) values. Results taken from MPD and ICC serve as the basis for identifying variability between imaging sequences. Though the primary focus of this study is within-subject reproducibility, *post-hoc* analyses were performed using t-tests to compare ICC values across ROIs for bilateral thickness, volume, and area, and MPD values (for each ROI) between the ASD and TDC groups to determine if any systematic between group differences in reproducibility exist. A general linear model was also conducted to compare global and regional thickness, volume, and area values between ASD and TDC, controlling for participant age, for both MPRAGE and MP2RAGE sequences to assess whether any between-group differences were reproducible across scans. False discovery rate was used to correct for multiple comparisons for between group analyses. Brain images were created using BrainNet Viewer (Xia et al., 2013).

3. Results

In the final sample of 163 male participants ages 14.7–57.9 (mean age = 27.28 ± 7.33 years), high consistency between MPRAGE and MP2RAGE scans for volume and surface area estimates were found (see Table 1). Average cortical thickness, total surface area, and whole brain volumetric measures all demonstrated good to excellent ICC (see Table 1 and Supplementary Material Fig. 2 (Bland-Altman plots)).

3.1. Regional volume

Volumetric reliability by Desikan-Killiany region is shown in Fig. 1 (ICC) and Fig. 2 (MPD). Most regions exhibited good (0.75-0.90) to excellent (>0.90) ICC values between MPRAGE and MP2RAGE derived bilateral brain volumes with the exception of entorhinal (ICC = 0.71), frontal (ICC = 0.60) and temporal (ICC = 0.60) polar cortices which had the lowest values and had the largest MPD in volume between scans (see Tables 1 and 2 in the Supplementary Material for regional ICC and MPD values). No bilateral ROIs

for volume had ICC values less than 0.60. A comparison of total brain segmented volume (excluding ventricles) resulted in a 0.99 ICC and 0.29% MPD demonstrating that global volume measurements are highly reliable between MPRAGE and MP2RAGE despite limited regional variation (see Fig. 3).

3.2. Surface area

The ICC values for bilateral surface area measurements were all equal to or greater than 0.75 when comparing MPRAGE to MP2RAGE scans apart from the insula (ICC = 0.65), frontal (ICC = 0.68) and temporal (ICC = 0.64) poles. No surface area ROIs presented ICC values below 0.64 (see Fig. 1 and Supplementary Material Table 1). These data suggest that when MRIs are taken with either imaging method and reconstructed using surface-based analysis, cortical surfaces closely resemble each other in most areas but several ROIs exhibit systematic morphometric variation. This notion is further supported by MPD values which show the absolute variation in each ROI between imaging methods. Mean MPD values for bilateral surface area (adding ROIs together) was 1.03% with the largest variance at 2.58% and the smallest at 0.50% reinforcing the trend that some regions of the brain are more difficult to replicate than others (see Fig. 2 and Supplementary Material Table 2).

3.3. Cortical thickness

Cortical thickness measurements exhibit higher generalized variability and appear to be less reproducible across scan types (see Figs. 1 and 2). Only four bilateral ROI exhibited greater than 0.75 ICC between MPRAGE and MP2RAGE (paracentral, parahippocampal, precentral, transverse temporal). The average ICC among all ROIs was 0.65. Cortical thickness was most variable (less than 0.50 ICC) in the bilateral frontal (ICC = 0.41) and temporal (ICC = 0.35) (see Supplementary Material Table 1). On average, MPD was only 1.17% across ROIs for cortical thickness comparing MPRAGE with MP2RAGE (see Supplementary Material Table 2).

3.4. MPRAGE vs MP2RAGE in diagnostic groups post hoc comparisons

When we compared ICC and MPD values between the ASD and TDC groups, no significant group differences were found for global or regional volumes, area or cortical thickness. We also examined ASD vs TDC differences in global and regional area, volume and thickness measures for each sequence controlling for participant age. o between-group differences were found for any of the global measures. Decreased area and volume in the right inferior temporal cortex and thinner bilateral insula was found in ASD in both sequences. A summary of all regional differences is provided as Supplementary Material (Supplementary Material Fig. 3).

In summary, MPRAGE and MP2RAGE are highly consistent for gray matter surface area and volume estimates but some regions demonstrate more systematic variability between imaging sequences. Cortical thickness measurements exhibit higher generalized variability and are less reliable overall (see Fig. 4).

4. Discussion

We find that reproducibility of volumetric and surface area gray matter measurements obtained from MPRAGE and MP2RAGE images show generally good to excellent reproducibility (mean ICC for surface area was 0.91 and 0.88 for volume) and showed only moderate to good reproducibility when more complex measurements such as cortical thickness were considered (mean ICC for cortical thickness was 0.65). However, we found spatial heterogeneity of reproducibility across the brain, with some regions showing much poorer reproducibility across sequence types.

There are potential sources of increased variance and poorer reproducibility in several of these regions. The inferior temporal cortex is in close proximity to the temporal bones, with high magnetic susceptibility adjacent to air-bone interfaces in the mastoid air cells, with the temporal bone constituting the densest bone in the body. Similarly, the frontal pole lies in a region of high susceptibility associated with artifacts from the frontal sinuses and dental hardware. These regions demonstrated the highest variability across surface area, volume, and thickness. Nine individuals in this study had permanent retainers, however their scans showed no visible artifacts upon inspection. A concern arose after the study that pooling data with ASD and TDC individuals would create a potential source of variation due to increased head movement in ASD individuals, but no group differences were found.

Nevertheless, it is overall reassuring that vety high ICC and low MPD between volumetric and surface area measurements are found across the brain, suggesting that both MPRAGE and MP2RAGE obtained on two different scanners and head coils can be incorporated into longitudinal analyses with relatively low risk of introducing regionally specific artifacts. These findings are consistent with studies showing high repeatability of FreeSurfer measurements for deep gray matter regions between MPRAGE and MP2RAGE images (Okubo, Okada, Yamamoto, et al.). These data are also consistent with findings of high repeatability of structural imaging metrics across segmentation algorithms (de Boer et al., 2010; Klauschen, Goldman, Barra, Meyer-Lindenberg, Lundervold), sites (Jovicich et al., 2006), scanner strength (Fujimoto et al., 2014), and computing platforms (Glatard et al., 2015). Additionally, significant differences in thickness reproducibility have been replicated when looking at reproducibility across different combinations of imaging sequence and head coil (Yan et al., 2020).

The inclusion of different scanner types accurately models real-world conditions that researchers often face when using multi-site or longitudinal datasets, however it is unknown how scanner hardware (Trio vs Prisma) and different head coils (12- vs 64-channel) used to collect the MPRAGE vs MP2RAGE sequences impacted the study findings. Although the FreeSurfer pipeline samples input to isotropic resolution, the MPRAGE and MP2RAGE acquisition sequences were acquired at different resolutions which could lead to interpolation artifacts. However, our results suggest that this did not affect volume or area estimates but may have affected FreeSurfer cortical thickness estimates. Finally, our age range consisted of individuals from adolescence to adulthood. It is unknown how regional volume, area and cortical thickness reliability estimates would differ in children or older adults.

5. Conclusions

Data acquired using MPRAGE (Siemens Trio; 12-channel head coil) is highly consistent with data acquired with MP2RAGE (Siemens Prisma; 64-channel head coil) but reasonable caution should be used when comparing structural differences in specific regions of the brain. Researchers pooling data from multiple sources should pay attention to scan type and scanner hardware when performing analysis and be aware of the margin of error stemming from scan type differences. Researchers should use caution when interpreting pooled database findings based on regions known to have low structural reproducibility, including posterior cingulate, medial prefrontal, and temporoparietal junction cortical regions. The findings in this project further emphasize the need for strict quality control measures and standardized datasets in national databases leveraging data pooled from multiple sources.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Carolyn King and Jubel Morgan for their contributions to this work.

Funding

Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under Award Numbers R01MH080826 and K08MH100609; data acquisition was supported by S10OD018482. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Alonso J, Pareto D, Alberich M, et al. , Dec 2020. Assessment of brain volumes obtained from MP-RAGE and MP2RAGE images, quantified using different segmentation methods. Magma 33 (6), 757–767. 10.1007/s10334-020-00854-4. [PubMed: 32468150]

de Boer R, Vrooman HA, Ikram MA, et al. , Jul 1 2010. Accuracy and reproducibility study of automatic MRI brain tissue segmentation methods. Neuroimage 51 (3), 1047–1056. 10.1016/j.neuroimage.2010.03.012. [PubMed: 20226258]

Dale AM, Fischl B, Sereno MI, Feb 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9 (2), 179–194. 10.1006/nimg.1998.0395. [PubMed: 9931268]

Desikan RS, Segonne F, Fischl B, et al. , Jul 1 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31 (3), 968–980. 10.1016/j.neuroimage.2006.01.021. [PubMed: 16530430]

van Erp TGM, Walton E, Hibar DP, et al. , Nov 1 2018. Cortical brain abnormalities in 4474 individuals with schizophrenia and 5098 control subjects via the enhancing neuro imaging genetics through meta analysis (ENIGMA) consortium. Biol. Psychiatr 84 (9), 644–654. 10.1016/j.biopsych.2018.04.023.

Eshaghi A, Marinescu RV, Young AL, et al. , Jun 1 2018. Progression of regional grey matter atrophy in multiple sclerosis. Brain 141 (6), 1665–1677. 10.1093/brain/awy088. [PubMed: 29741648]

Fischl B, Dale AM, Sep 26 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U.S.A 97 (20), 11050–11055. 10.1073/pnas.200033797. [PubMed: 10984517]

Fischl B, van der Kouwe A, Destrieux C, et al. , Jan 2004. Automatically parcellating the human cerebral cortex. Cerebr. Cortex 14 (1), 11–22. 10.1093/cercor/bhg087.

Fischl B, Salat DH, Busa E, et al. , Jan 31 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33 (3), 341–355. 10.1016/s0896-6273(02)00569-x. [PubMed: 11832223]

Fischl B, Sereno MI, Dale AM, Feb 1999. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. Neuroimage 9 (2), 195–207. 10.1006/nimg.1998.0396. [PubMed: 9931269]

Frisoni GB, Fox NC, Jack CR Jr., Scheltens P, Thompson PM, Feb 2010. The clinical use of structural MRI in Alzheimer disease. Nat. Rev. Neurol 6 (2), 67–77. 10.1038/nrneurol.2009.215. [PubMed: 20139996]

Fujimoto K, Polimeni JR, van der Kouwe AJ, et al. , Apr 15 2014. Quantitative comparison of cortical surface reconstructions from MP2RAGE and multi-echo MPRAGE data at 3 and 7 T. Neuroimage 90, 60–73. 10.1016/j.neuroimage.2013.12.012. [PubMed: 24345388]

Glatard T, Lewis LB, Ferreira da Silva R, et al. , 2015. Reproducibility of neuroimaging analyses across operating systems. Front. Neuroinf 9, 12. 10.3389/fninf.2015.00012.

Gronenschild EH, Habets P, Jacobs HI, et al. , 2012. The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements. PLoS One 7 (6), e38234. 10.1371/journal.pone.0038234. [PubMed: 22675527]

Han X, Jovicich J, Salat D, et al. , Aug 1 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. Neuroimage 32 (1), 180–194. 10.1016/j.neuroimage.2006.02.051. [PubMed: 16651008]

Jack CR Jr., Bernstein MA, Fox NC, et al. , Apr 2008. The alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imag 27 (4), 685–691. 10.1002/jmri.21049.

Jovicich J, Czanner S, Greve D, et al. , Apr 1 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. Neuroimage 30 (2), 436–443. 10.1016/j.neuroimage.2005.09.046. [PubMed: 16300968]

Kecskemeti S, Samsonov A, Hurley SA, Dean DC, Field A, Alexander AL, Mar 2016. MPnRAGE: a technique to simultaneously acquire hundreds of differently contrasted MPRAGE images with applications to quantitative T1 mapping. Magn. Reson. Med 75 (3), 1040–1053. 10.1002/mrm.25674. [PubMed: 25885265]

Klauschen F, Goldman A, Barra V, Meyer-Lindenberg A, Lundervold A, Apr 2009. Evaluation of automated brain MR image segmentation and volumetry methods. Hum. Brain Mapp 30 (4), 1310–1327. 10.1002/hbm.20599. [PubMed: 18537111]

Koshiyama D, Fukunaga M, Okada N, et al. , Jan 10 2018. Subcortical association with memory performance in schizophrenia: a structural magnetic resonance imaging study. Transl. Psychiatry 8 (1), 20. 10.1038/s41398-017-0069-3. [PubMed: 29317603]

Lange N, Travers BG, Bigler ED, et al. , Feb 2015. Longitudinal volumetric brain changes in autism spectrum disorder ages 6-35 years. Autism Res. 8 (1), 82–93. 10.1002/aur.1427. [PubMed: 25381736]

Marques JP, Kober T, Krueger G, van der Zwaag W, Van de Moortele PF, Gruetter R, Jan 15 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage 49 (2), 1271–1281. 10.1016/j.neuroimage.2009.10.002. [PubMed: 19819338]

Di Martino A, 2012. The Autism Brain Imaging Data Exchange (ABIDE) Consortium: Open Sharing of Autism Resting State fMRI.

McGraw KO, Wong SP, 1996. Forming inferences about some intraclass correlation coefficients. Psychol. Methods 1 (1), 30–46. 10.1037/1082-989X.1.1.30.

Mugler JP 3rd, Brookeman JR, Jul 1990. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). Magn. Reson. Med 15 (1), 152–157. 10.1002/mrm.1910150117. [PubMed: 2374495]

Okubo G, Okada T, Yamamoto A, et al. , Jan 2016. MP2RAGE for deep gray matter measurement of the brain: a comparative study with MPRAGE. J. Magn. Reson. Imag 43 (1), 55–62. 10.1002/jmri.24960.

Reuter M, Schmansky NJ, Rosas HD, Fischl B, Jul 16 2012. Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage 61 (4), 1402–1418. 10.1016/j.neuroimage.2012.02.084. [PubMed: 22430496]

Salarian A, 2019. Intraclass correlation coefficient (ICC). https://www.mathworks.com/matlabcentral/fileexchange/22099-intraclass-correlation-coefficient-icc. (Accessed 3 September 2019).

Xia M, Wang J, He Y, 2013. BrainNet Viewer: a network visualization tool for human brain connectomics. PLoS One 8 (7), e68910. 10.1371/journal.pone.0068910. [PubMed: 23861951]

Yan S, Qian T, Marechal B, et al. , Jan 2020. Test-retest variability of brain morphometry analysis: an investigation of sequence and coil effects. Ann. Transl. Med 8 (1), 12. 10.21037/atm.2019.11.149. [PubMed: 32055603]

Zielinski BA, Prigge MB, Nielsen JA, et al. , Jun 2014. Longitudinal changes in cortical thickness in autism and typical development. Brain 137 (Pt 6), 1799–1812. 10.1093/brain/awu083. [PubMed: 24755274]

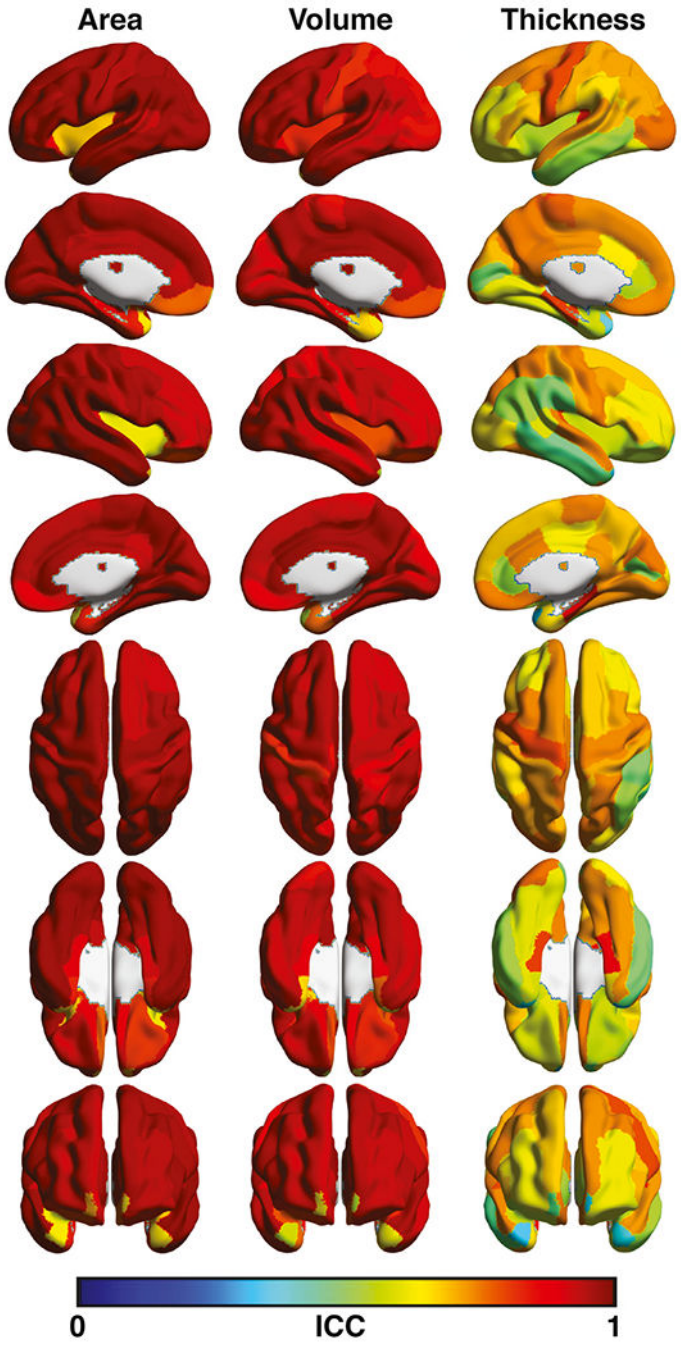

Fig. 1.

Intraclass correlation coefficients (ICC) overlaid across the Desikan-Killiany atlas demonstrate which regions of the brain are most susceptible to structural variability between MPRAGE and MP2RAGE sequences.

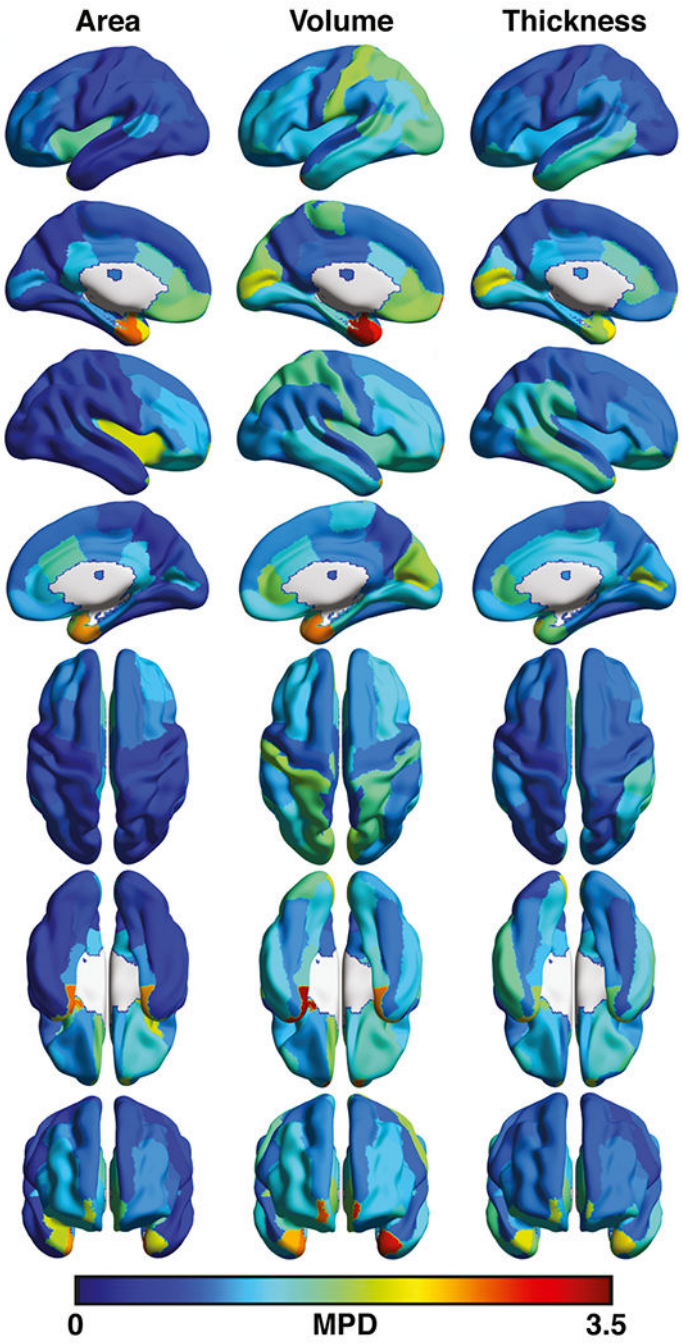

Fig. 2.

Mean percent difference (MPD) values overlaid across the Desikan-Killiany atlas demonstrate which regions of the brain are most susceptible to structural variability between MPRAGE and MP2RAGE sequences.

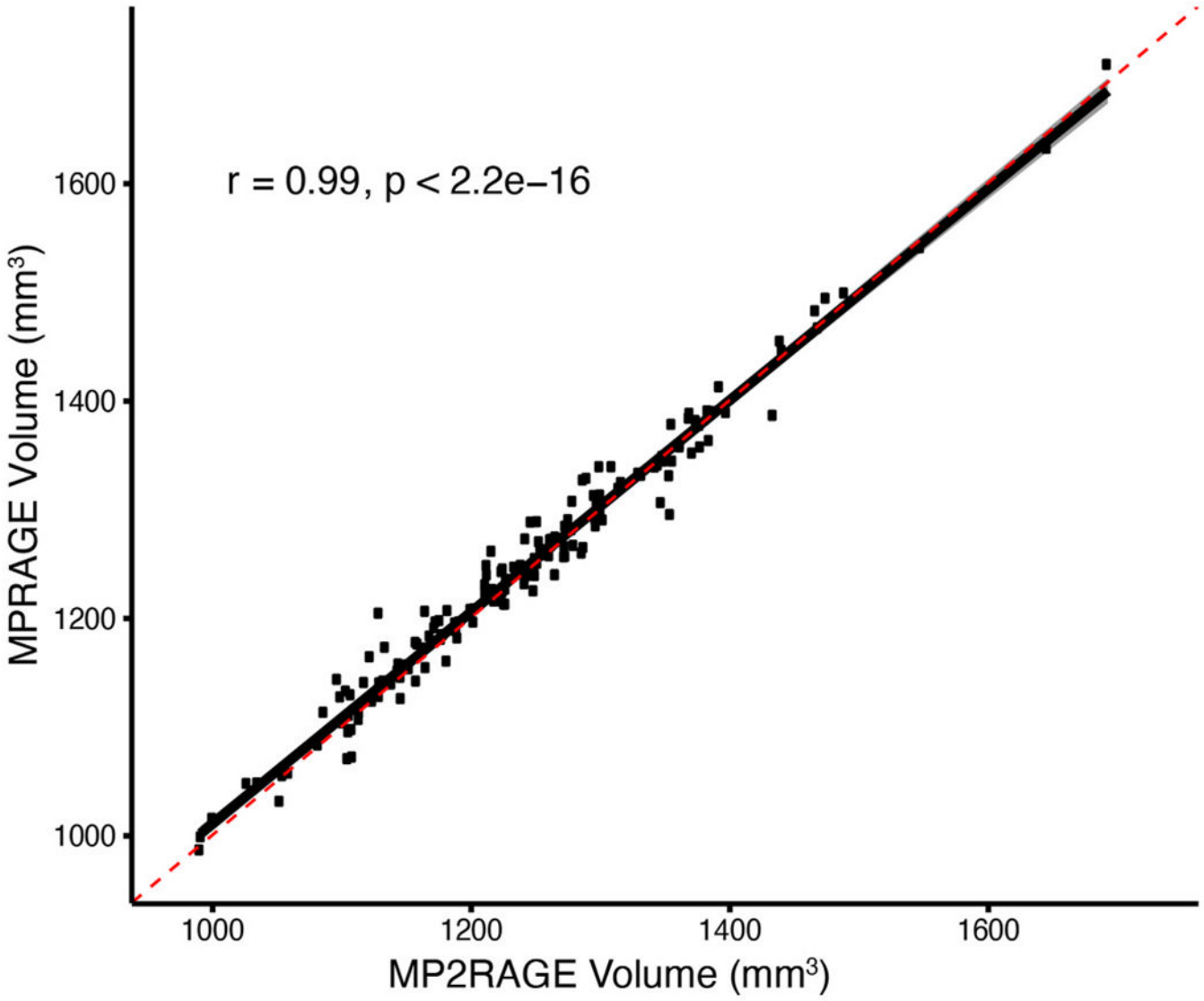

Fig. 3.
Total brain volumes derived from MPRAGE and MP2RAGE show high interscan reliability.

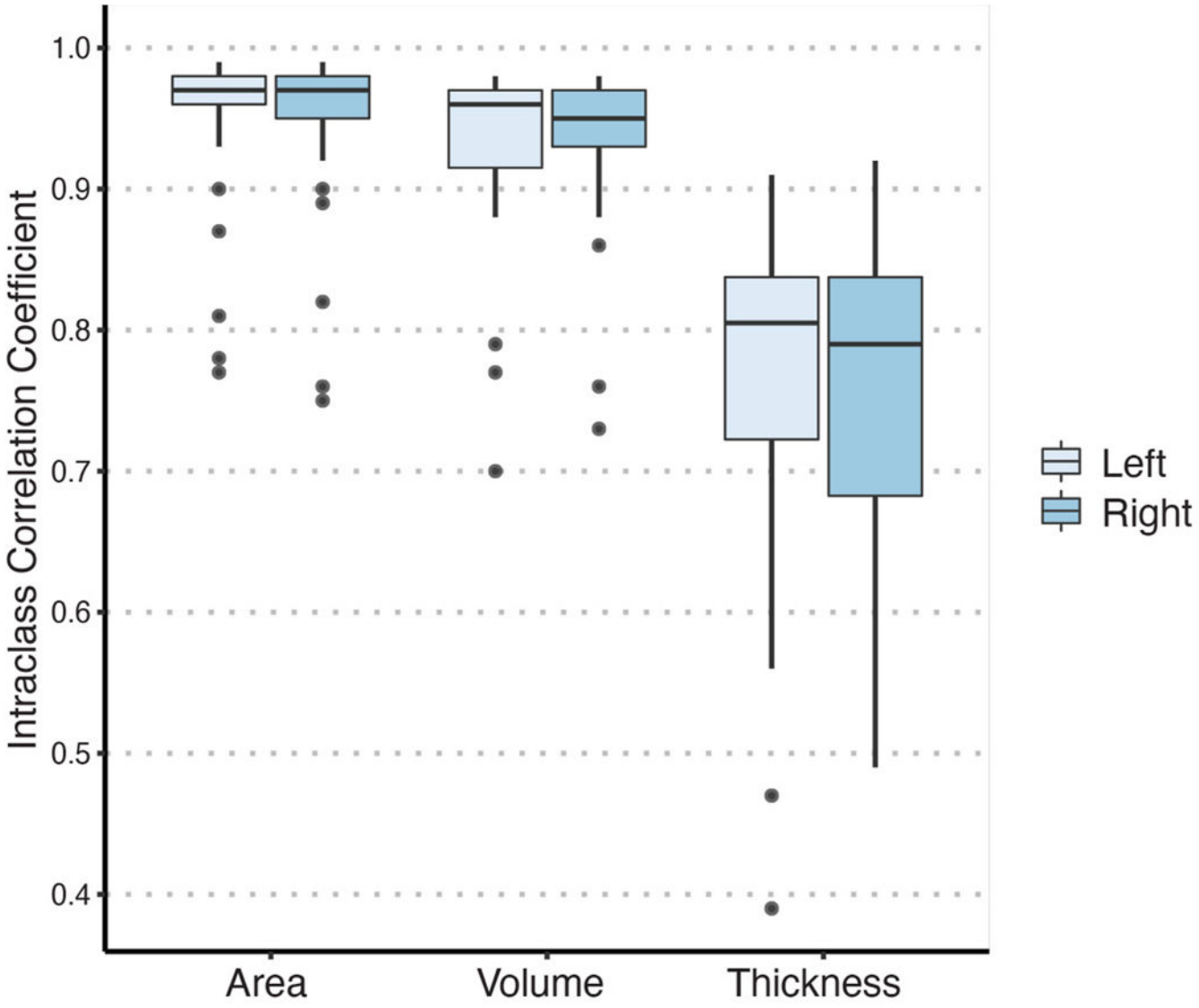

Fig. 4.
Distribution of intraclass correlation coefficient values across ROIs between MPRAGE and MP2RAGE.

Table 1

Intraclass correlation coefficient and mean percent difference values for brain measures acquired using MPRAGE and MP2RAGE sequences.

Measure of Analysis	MPRAGE		MP2RAGE		Average	
	Mean	SD	Mean	SD	ICC	MPD
Avg. Cortical Thickness (mm)	2.54	0.09	2.50	0.12	0.75	0.65
Total Surface Area (mm^2)	187502.32	15409.92	190828.85	15472.83	0.97	0.48
Gray Matter Volume (mm^3)	717838.03	63914.65	710648.37	66530.46	0.95	0.52
White Matter Volume (mm^3)	497570.52	57849.77	505041.17	59016.94	0.99	0.42
Total CSF Volume (mm^3)	19394.00	8970.36	23311.36	9632.81	0.95	4.97
Total Brain Volume (mm^3)	1244921.10	113663.97	1239933.63	115206.60	0.99	0.29

Abbreviations: CSF, cerebrospinal fluid; ICC, intraclass correlation coefficient; MPD, mean percent difference; SD, standard deviation.