NOMIS: Quantifying morphometric deviation from normality over the lifetime in the adult human brain

Olivier Potvin[1], PhD, Louis Dieumegarde[1], BSc, and Simon Duchesne, PhD[12] for the Alzheimer's Disease Neuroimaging Initiative* and CIMA-Q**

[1] CERVO Brain Research Centre, 2601, de la Canardière, Québec, Canada, G1J 2G3

[2] Département de radiologie, Faculté de médecine, Université Laval, 1050, avenue de la Médecine, Québec, Canada, G1V 0A6

Key words: magnetic resonance imaging, atrophy, morphometry, normality, aging, sex.

Correspondence:

Simon Duchesne, PhD
CERVO Brain Research Centre
F-3582, 2601, de la Canardière, Québec, Canada, G1J 2G3
Phone: 418 663-5000 ext.4777
Fax: 418 663-9540
simon.duchesne@fmed.ulaval.ca

*Part of the data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (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

** Part of the data used in this article were obtained from the Consortium pour l'identification précoce de la maladie Alzheimer - Québec (CIMA-Q; cima-q.ca). As such, the investigators within the CIMA-Q contributed to the design, the implementation, the acquisition of clinical, cognitive, and neuroimaging data and biological samples. A list of the CIMA-Q investigators is available on www.cima-q.ca.

Abstract

We present NOMIS (https://github.com/medicslab/NOMIS), a comprehensive open MRI tool to assess morphometric deviation from normality in the adult human brain. Based on MR anatomical images from 6,909 cognitively healthy individuals aged 18-100 years, we modeled 1,344 measures computed using the open access *FreeSurfer* pipeline, taking into account personal characteristics (age, sex, head size), scanner characteristics (manufacturer and magnetic field strength), and image quality, providing expected values for any new individual. Then, for each measure, the NOMIS tool was built to generate Z-score effect sizes denoting the extent of deviation from the normative sample. Depending on the user need, NOMIS offers four versions of Z-score adjusted on different sets of variables. While all versions take into account head size, image quality and scanner characteristics, they can also incorporate age and/or sex, thereby facilitating multi-site neuromorphometric research across adulthood.

Introduction

Despite the popularity of magnetic resonance imaging (MRI) to examine abnormalities in brain morphometry, tools quantifying normality are lacking. While age, sex and intracranial volume are well-known to influence brain volume and shape[1, 2] the determination of whether an individual's brain region measurements are within normality faces multiple major challenges such as the lack of normative data across appropriate age groups, the influence of the MRI processing pipeline, the variety in neuroanatomical atlases used for parcellation and the quality of the image itself[3, 4]. We made previous attempts[5-8] to produce such normative data in adulthood based on *FreeSurfer*, an open-access and fully automated segmentation software (http://freesurfer.net), for two specific brain atlases, namely Desikan-Killiany[9] (DK) and Desikan-Killiany-Tourville[10] (DKT). This initial foray allowed for the quantification of the extent of deviation from normality for a given individual, according to personal characteristics such as age, sex and estimated intracranial volume (eTIV), while controlling for scanner magnetic field strength (MFS) and scanner manufacturer (OEM). Although this work has already gathered more than a hundred citations in the last three years, several researchers solicited the expansion of the norms to include other brain regions, as well as different atlases; as well as the production of variants, for example, of only adjusting for head size and scanner characteristics. For the latter, it became clear to us that we also needed to control for image quality. Further, we recognized the need to increase the size of our normative sample to ensure better representation of middle age.

Leveraging this prior work, we offer a comprehensive tool called NOMIS (NOrmative Morphometry Image Statistics; https://github.com/medicslab/NOMIS). NOMIS can be used for

new adult individuals, healthy or otherwise. Using this individual's T1-weighted MRI, processed via the *FreeSurfer* 6.0 toolkit, one can derive Z-score effect sizes denoting the extent of deviation from the normative sample according to the individual's characteristics (age, sex, and eTIV), while taking into account scanner information (MFS, OEM) and now voxel size (resolution) and image quality (contrast-to-noise ratio (CNR) and holes in surface reconstruction). Our model takes into account 1,344 brain measures generated by *FreeSurfer* on 6,909 healthy individuals aged 18 to 100 years (mean ± sd: 55.0 ± 20.0; 56.8% female). The normative data includes as before the DK[9] and DKT[10] atlases, as well as the Destrieux (a2009s)[11] neocortical atlas; neocortical pial and white surface areas, volumes and thicknesses; *FreeSurfer*'s default subcortical atlas[12], hippocampal subfields, brainstem subregions; its ex vivo-based labeling protocol atlas[13]; and the subcortical white matter parcellation according to the adjacent neocortical areas. Furthermore, to fulfill specific needs from researchers, we propose four versions of Z-score adjusted on different sets of variables. While all versions are adjusted for head size, image quality and scanner characteristics, the full version includes both age and sex whereas the three other versions are without age, without sex and without age and sex. Thus, a research group working on aging aiming at removing the variance of hippocampal volumes due to head size, sex, scanner, and image quality could use the version without age, which preserves the variance due to aging.

Materials and methods

Normative sample

The norms are based on a cross-sectional sample of 6,909 (initial sample: 7,399) cognitively healthy individuals aged 18 to 100 years, (mean ±sd; 55.0 ±20.0; 56.8% female), gathered from

91 27 different datasets (Table 1; Acknowledgments). Scans were acquired from one of the three
92 leading OEM (e.g. Siemens Healthcare (Erlangen, Germany); Philips Medical Systems (Best,
93 Netherlands); or GE Healthcare (Milwaukee, WI)) at MFS of either 1.5 or 3 Tesla. For each dataset,
94 approval from the local ethics board and informed consent of the participants were obtained.

95 **Table 1. Datasets included in the normative sample**

Dataset	n
Autism Brain Imaging Data Exchange (ABIDE)[14]	183
Alzheimer's Disease Neuroimaging Initiative (ADNI)[15]	672
Australian Imaging Biomarkers and Lifestyle flagship study of ageing (AIBL)[16]	157
Berlin Mind and Brain (Margulies, Villringer) CoRR sample (BMB)[17, 18]	50
Cambridge Centre for Ageing and Neuroscience (CamCAN)[19, 20]	630
Center of Biomedical Research Excellence (COBRE)[21]	70
Cleveland CCF[22]	30
Consortium for the Early Identification of Alzheimer's Disease (CIMA-Q)[23]	29
Dallas Lifespan Brain Study (DLBS)[24]	304
FIND lab sample (FIND) Lab[25]	13
Functional Biomedical Informatics Research Network (FBIRN)[26]	33
Lifespan Human Connectome Project in Aging (HCP-Aging)[27]	612
International Consortium for Brain Mapping (ICBM) - MNI[28]	147
Information eXtraction from Images (IXI)[29]	554
F.M. Kirby Research Center neuroimaging reproducibility data (KIRBY-21)[30]	20
Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD)[31]	21
National Alzheimer's Coordinating Center (NACC)[32]	1562
National Database for Autism Research (NDAR)[33]	56
Nathan Kline Institute Rockland (NKI-RS)[34]	138
Nathan Kline Institute Rockland (NKI-RS) Enhanced[34]	436
Open Access Series of Imaging Studies (OASIS)[35]	288
POWER Neuroimage sample (POWER)[36]	26
Parkinson's Progression Markers Initiative (PPMI)[37]	158
Southwest University Adult Lifespan Dataset (SALD)[38]	490
University of Wisconsin (Birn, Prabhakaran, Meyerand) CoRR sample (UWM)[17]	25
Wayne State EF Dataset[39]	108
Yale Low-Resolution Controls Dataset[40]	97
Total	6909

Among the datasets are the Alzheimer's Disease Neuroimaging Initiative (ADNI), the Australian Imaging, Biomarkers and Lifestyle study of aging (AIBL) and the Consortium pour l'identification précoce de la maladie Alzheimer - Québec (CIMA-Q) datasets. The ADNI (adni.loni.usc.edu) was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. (www.adni-info.org). The AIBL data was collected by the AIBL study group and AIBL study methodology has been reported previously by Ellis et al. (2009). For each dataset, approval from the local ethics board and informed consent of the participants were obtained. Founded in 2013 with a $2,500,000 grant from the Fonds d'Innovation Pfizer - Fond de Recherche Québec – Santé sur la maladie d'Alzheimer et les maladies apparentées, the main objective of CIMA-Q is to build a cohort of participants characterized in terms of cognition, neuroimaging and clinical outcomes in order to acquire biological samples allowing (1) to establish early diagnoses of Alzheimer's disease, (2) to provide a well characterized cohort and (3) to identify new therapeutic targets. The principal investigator and director of CIMA-Q is Dr Sylvie Belleville from the Centre de recherche de l'Institut universitaire de gériatrie de Montréal, CIUSSS Centre-sud-de-l'Île-de-Montréal. CIMA-Q represent a common effort of several researchers from Québec affiliated to Université Laval, Université McGill, Université de Montréal, et Université de Sherbrooke. CIMA-Q recruited 350 cognitively healthy participants, with subjective cognitive impairment, mild cognitive impairment, or Alzheimer's disease, between 2013–2016. Volunteers were recruited from memory clinics, through advertisements posted in the community and amongst participants from the NuAge population study.

From all the samples mentioned, only cognitively healthy (control) participants were included. For the Nathan Kline Institute samples, which were projects recruiting in the general population, we excluded participants with history of schizophrenia or other psychotic disorders, bipolar disorders, major depressive disorders (recurrent), posttraumatic stress disorder, substance abuse/dependence disorders, neurodegenerative and neurological disorders, head injury with loss of consciousness/amnesia, and lead poisoning. Moreover, for the Parkinson's Progression Markers Initiative dataset, we excluded participants with a Geriatric Depression Scale[41] score of more than 5.

Brain segmentation

Brain segmentation was conducted using *FreeSurfer* version 6.0, a widely used and freely available automated processing pipeline that quantifies brain anatomy (http://freesurfer.net). All raw T1-weighted images were processed using the "recon-all -all" *FreeSurfer* command with the fully-automated directive parameters (no manual intervention or expert flag options) on the CBRAIN platform[42]. Normative data were computed for volumes, neocortical thicknesses and white and pial surfaces areas for all atlases comprised in *FreeSurfer* 6.0: the default subcortical atlas[12] (aseg.stats), the Desikan-Killiany atlas[9] (DK, aparc.stats file), the Desikan-Killiany-Tourville atlas[10] (DKT, aparc.DKT.stats file), the Destrieux atlas[11] (aparc.a2009s.stats file), the ex vivo atlas,[43] including entorhinal and perirhinal cortices, the brainstem sub-regions atlas[44], the Brodmann area maps which includes somatosensory areas, several motor and visual areas, as well as the hippocampal subfields atlas[45].

The technical details of *FreeSurfer*'s procedures are described in prior publications. Briefly, this processing includes motion correction, removal of non-brain tissue using a hybrid

139 watershed/surface deformation procedure, automated Talairach transformation, intensity
140 normalization, tessellation of the gray matter white matter boundary, automated topology
141 correction, and surface deformation following intensity gradients to optimally place the
142 gray/white and gray/cerebrospinal fluid borders at the location where the greatest shift in
143 intensity defines the transition to the other tissue class. Once the cortical models are complete,
144 a number of deformable procedures can be performed for further data processing and analysis
145 including surface inflation, registration to a spherical atlas which is based on individual cortical
146 folding patterns to match cortical geometry across subjects and parcellation of the cerebral
147 cortex into units with respect to gyral and sulcal structure. This method uses both intensity and
148 continuity information from the entire three-dimensional MRI volume in segmentation and
149 deformation procedures to produce representations of cortical thickness, calculated as the
150 closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the
151 tessellated surface. The maps are created using spatial intensity gradients across tissue classes
152 and are therefore not simply reliant on absolute signal intensity. Procedures for the
153 measurement of cortical thickness have been validated against histological analysis [46] and
154 manual measurements[47, 48]. Estimated intracranial volume[49] was taken from the aseg.stats
155 *FreeSurfer* output file. We added the total ventricle volume (labeled as "ventricles") using the
156 sum of all ventricles and the corpus callosum (labeled as "cc") using the sum of all corpus
157 callosum segments.

158 **Quality control and sample selection**

159 A flow chart detailing the final analysis sample is shown in Fig 1. From an initial pool of 7,399
160 MRIs, nine images failed the *FreeSurfer* pipeline. Following processing, each of the remaining

7,390 brain segmentations was visually inspected through at least 20 evenly distributed coronal sections. After quality control, 445 images (6.0%) were removed from further analyses due to segmentation problems, the main reason being that parts of the brain were not completely segmented (e.g. temporal and occipital poles). During visual inspection, 26 brains were found to have clear significant brain lesions and were excluded. In addition to visual inspection, we excluded participants if at least one of the 1,344 brain region measures was missing (n=10). In fine, the analysis sample was composed of 6,909 individual MRIs.

Fig 1. Flow chart of the images.

Training, validation and test sample

We randomly selected 10% of the whole sample (n=691) to test the models in an independent sample (age: 55.1 ±20.1, range 18-100; 58.5% female). This test sample was not used to build the models predicting normative values. The remaining 90% was used as training sample (age: 54.9 ±20.0, range 18-100; 56.7% female) to build and validate the models. Leave-10%-out cross-validation was used to validate the model in the training sample.

Clinical samples evaluations

We evaluated the usefulness of normative values using clinical samples of individuals with schizophrenia (n=72; Age: 38.2 ±13.9, range 18-65; 19% female) from the COBRE dataset, as well as participants with clinically ascertained mild Alzheimer's disease (n=157 Age: 74.8 ±8.1, range 55-90; 43% female) from the baseline ADNI-2 dataset.

Scanner-related predictors

Scanner-related predictors included manufacturer (OEM), magnetic field strength (MFS), and voxel size (resolution).

Image quality predictors

Image quality can have an effect on brain segmentation quality[50, 51]. We therefore included in the prediction models contrast-to-noise ratio (CNR) for each area as well as the total number of holes in surface reconstruction prior to fixing, since this measure is correlated with visual assessment of brain segmentation[50]. Defect holes – topological errors in the cortical surface reconstructions – were extracted from the aseg.stats *FreeSurfer* output file and CNR was assessed after *FreeSurfer* preprocessing using the brain.mgz file with gray matter (GM) and cerebral white matter (WM) voxel intensities for each area with the following formula:

$$CNR = \frac{(GM\ mean\ -\ WM\ mean\ )^2}{(GM\ variance\ +\ WM\ variance)}$$

Regression models and statistical analyses

Linear regression models predicting each measure were built using age, sex, eTIV, MFS, OEM, voxel size, surface defect holes and the CNR of the region as predictors. To obtain normal distributions, surface holes and ventricles (except the 4th) were log transformed. For ventricles and white matter regions, CNR of the total brain gray matter was used while for the brainstem subregions and hippocampal subfields, CNR from the whole brainstem and whole hippocampus were used, respectively. Quadratic and cubic terms for age, eTIV, and surface holes were included, as well as the following interactions: age X sex, age X eTIV, sex X eTIV, eTIV X MFS and MFS X OEM. Feature selection was conducted with a 10-fold cross-validation[52] backward

elimination procedure, retaining the model with the subset of predictors that produced the lowest predicted residual sum of squares. For each selected final model, the fit of the data was assessed using R^2 coefficient of determination:

$$R^2 = 1 - \frac{\sum_i (Y_i - f_i)^2}{\sum_i (Y_i - \hat{Y})^2}$$

where the numerator is the residual sum of squares (*Y* is the value of the variable to be predicted and *f* is the predicted value) and the denominator is the total sum of squares ($\hat{Y}$ is the mean). To assess the unique contribution of each predictor, we used the lmg metric in the *R* package[53] relaimpo[54]. This metric is a R^2 partitioned by averaging sequential sums of squares over all orderings of the predictors.

For each brain measure, in order to exclude potential abnormalities, outliers with Z scores higher than 3.29 ($p < .001$) were removed to compute the statistical model. This was done in proportion to eTIV for volumes and surfaces and on raw values for cortical thicknesses. The number of outliers was below 1% for all regions (mean ±sd of all atlases: 0.45% ±0.10%) except the right long insular gyrus and central sulcus of the insula white surface (1.1%) and pericallosal sulcus volume (1.1%) of the Destrieux atlas. Detailed results can be found in the supplementary material. Brain figures were made using the ggseg R package[55].

The models were verified by examining the difference between CV-10 R^2 of the training sample and R^2 of the independent test sample of healthy controls. We then examined the validity of the normative values to show expected patterns of normality deviations using the Z score effect sizes in the validation samples of healthy individuals and of individuals with AD and SZ.

$$Z_{OP} = \frac{Y_o - \hat{Y}}{RMSE}$$

Z score effect sizes (Z_{OP}) were obtained by subtracting the *Predicted value* ($\hat{Y}$) from the *Observed value* (Y_o) divided by the root mean square error (RMSE) of the model predicting the value [56].

Results

Total variance explained by the models

The cross-validation 10-fold (CV-10) R^2 for the overall explanatory variance of the models ranged between 0.02 to 0.84, with a mean ± sd of 0.37 ± 0.10. The highest R^2 were observed in the largest regions (total brain volume 0.84, neocortex volume 0.82, pial surface 0.80, and white surface 0.79). The inferior occipital gyrus and sulcus (0.02) and the anterior transverse collateral sulcus thicknesses in the Destrieux atlas (0.02), as well as the DK and DKT left parahippocampal thickness (0.02) had the lowest R^2. As examples, figures in this report display subcortical and DK neocortical atlases results. Full detailed results for all atlases are provided as supplementary information. Fig 2 illustrates the total R^2 for neocortical volumes and thicknesses of the DK atlas parcellation, as well as subcortical volumes. As shown, a higher amount of variance was generally explained for volumes compared to cortical thicknesses.

Independent test

The models predicting normative values were tested in an independent, healthy adults randomly chosen 10% sample. Nearly all models showed equivalent or higher R^2 on the test set than on the training set by CV-10 (difference for all atlases: 0.01 ±0.016). The lowest test differences were in the Destrieux atlas where 37 measures out of 592 were below -0.05, the worse being the fronto-marginal gyrus (of Wernicke) and sulcus (0.11), the superior occipital gyrus (0.10) and the superior temporal sulcus (0.10) pial surface areas. Fig 2 displays the R^2 difference between training and test sets for the DK atlas. Few minimal worse test values were

observed for some cortical and subcortical volumes while all cortical thicknesses and most other volumes had equivalent or better R^2 prediction values.

Fig 2. Top: R^2 from 10-fold cross-validation (10-fold CV) for cortical volumes and thicknesses from the DK atlas and subcortical volumes. Bottom: Difference between R^2 from the training set (10-fold CV) and the independent test set. Worse prediction from the test set are shown in green while better prediction are grayed.

Variance explained by each predictor

Figs 3 and 4 show variances due to biological and MRI factors. As expected, effects differ highly from one region to another, however, globally, age and eTIV had the largest effects on volumes, while age was the essential factor on neocortical thickness. Scanner, as well as image quality had statistically significant, but smaller effects on morphometric measures. Fig 5 illustrates the effect of age and sex on the four different NOMIS Z scores versions on the independent healthy test set. While the age effect is clearly apparent on the sex and the version without age and sex, it is null on the age and sex and age Z scores versions. The sex effect is also affected, but is relatively small.

Fig 3. Effects of age, sex, and estimated total intra-cranial volume (eTIV) on cortical volumes and thicknesses from the DK atlas and subcortical volumes.

Fig 4. Effects of scanner and image quality on cortical volumes and thicknesses from the DK atlas and subcortical volumes.

Fig 5. Age and sex effects on the left cortical thickness across the four NOMIS Z scores alternatives.

Clinical validation

We validated the normative values in individuals with clinically ascertained mild Alzheimer's disease and schizophrenia, which showed expected patterns of deviations from otherwise cognitively/behaviorally healthy individuals (Fig 6). In the Alzheimer's disease group, deviations from normality covered the frontal, temporal and parietal cortices with enlarged ventricles, but were especially more pronounced in the hippocampus and entorhinal cortex. In schizophrenia, atrophy was widespread to nearly all of the cortex.

Fig 6. Mean normative Z scores on cortical volumes and thicknesses from the DK atlas and subcortical volumes of participants with mild Alzheimer's disease and with schizophrenia.

Discussion

NOMIS strengths and limits

Prior normative data[5-7] were relatively limited in terms of atlases and sample size. With nearly seven thousand participants and 1,344 brain measures, NOMIS offers a comprehensive neuromorphometric normative tool based on a very large sample. In addition, an innovation of NOMIS is its flexibility. Depending on the user need, it has four versions of Z-score adjusted on different sets of variables. All versions include head size, image quality and scanner characteristics, but can also take into account age and/or sex or without age and sex. Therefore,

research groups looking for traditional norms, as well as others wanting to lower the variance due to head size and image quality/scanner characteristics while preserving age and/or sex variances can take advantage of NOMIS. Another strength of NOMIS is that the normative values were created on a various amalgam of cognitively healthy participants from multiple countries, with data acquired from a wide variety of MRI scanners and image quality, maximizing its generalizability. A novelty to prior existing normative data, is the addition of the image quality impact on the morphometry measures. Figure 4 shows that its effect is not trivial on cortical volume and thickness. Thus, our new normative data should help to remove some noise due to image quality.

Despite these strengths, users should keep in mind that before using NOMIS, it is mandatory to verify *FreeSurfer* segmentations and that while it will remove parts of variance due to the scanner and image quality, it won't correct for segmentation errors or image artefacts. Moreover, the normative sample, comprised essentially of research volunteers in academic-led environments, was recruited using a non-probability sampling method and may not be representative of the targeted population by the user.

Using NOMIS

The NOMIS tool is a user-friendly automated script in Python, freely accessible (https://github.com/medicslab/NOMIS). Users only need to pre-process their images with *FreeSurfer* 6.0 using automated directive parameters, then specify the individuals' characteristics to the script, which will automatically compute Z-scores based on the *FreeSurfer* output. One can choose the version of the Z-score by including in the csv file, only the variables that needs to be adjusted and the script automatically selects the appropriate version of predictors. The predictive

models and all statistical parameters are provided along with the script. We anticipate that this tool will be of broad interest to the neuroscientific community.

316 # Financial Disclosure Statement

317 OP and LD are supported by grants from the Canadian Consortium on Neurodegeneration in
318 Aging (CCNA) and the Canadian Institutes of Health Research (#IC119923) for which SD is a co-
319 investigator and a primary investigator, respectively. CCNA is supported with funding from
320 several partners including the Alzheimer Society of Canada, Sanofi, and Women's Brain Health
321 Initiative.

322

323 # Acknowledgments

324 This study comprises multiple samples of healthy individuals. We wish to thank all principal
325 investigators who collected these datasets and agreed to let them accessible.

326

327 Autism Brain Imaging Data Exchange (ABIDE): Primary support for the work by Adriana Di
328 Martino was provided by the NIMH (K23MH087770) and the Leon Levy Foundation. Primary
329 support for the work by Michael P. Milham and the INDI team was provided by gifts from
330 Joseph P. Healy and the Stavros Niarchos Foundation to the Child Mind Institute, as well as by
331 an NIMH award to MPM (R03MH096321). http://fcon_1000.projects.nitrc.org/indi/abide/

332

333 Alzheimer's Disease Neuroimaging Initiative (ADNI): The investigators within the ADNI
334 contributed to the design and implementation of ADNI and/or provided data but did not
335 participate in analysis or writing of this report. A complete listing of ADNI investigators can be
336 found at: http://adni.loni.usc.edu/wp-

content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. ADNI 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 (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.

http://adni.loni.usc.edu/

358 Australian Imaging Biomarkers and Lifestyle flagship study of ageing (AIBL): Part of the data
359 used in this study was obtained from the Australian Imaging Biomarkers and Lifestyle flagship
360 study of ageing (AIBL) funded by the Commonwealth Scientific and Industrial Research
361 Organisation (CSIRO) which was made available at the ADNI database
362 (www.loni.usc.edu/ADNI). The AIBL researchers contributed data but did not participate in
363 analysis or writing of this report. AIBL researchers are listed at www.aibl.csiro.au
364

365 Berlin Mind and Brain (Margulies, Villringer) CoRR sample (BMB). Zuo, X.N., et al. (2014). An
366 open science resource for establishing reliability and reproducibility in functional connectomics.
367 *Scientific data, 1*, 140049. doi: 10.1038/sdata.2014.49.
368 http://fcon_1000.projects.nitrc.org/indi/CoRR/html/bmb_1.html
369

370 Cambridge Centre for Ageing and Neuroscience (CamCAN): CamCAN funding was provided by the
371 UK Biotechnology and Biological Sciences Research Council (grant number BB/H008217/1),
372 together with support from the UK Medical Research Council and University of Cambridge, UK.
373 http://www.mrc-cbu.cam.ac.uk/datasets/camcan/
374

375 Center of Biomedical Research Excellence (COBRE): The imaging data and phenotypic
376 information was collected and shared by the Mind Research Network and the University of New
377 Mexico funded by a National Institute of Health COBRE: 1P20RR021938-01A2.
378 http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html
379

380 Cleveland Clinic (Cleveland CCF): Funded by the National Multiple Sclerosis Society.
381 http://fcon_1000.projects.nitrc.org/indi/retro/ClevelandCCF.html
382
383 Consortium for the Early Identification of Alzheimer's Disease (CIMA-Q): Part of the data used in
384 this article were obtained from the Consortium pour l'identification précoce de la maladie
385 Alzheimer - Québec (CIMA-Q). As such, the investigators within the CIMA-Q contributed to the
386 design, the implementation, the acquisition of clinical, cognitive, and neuroimaging data and
387 biological samples. A list of the CIMA-Q investigators is available on cima-q.ca.
388
389 Dallas Lifespan Brain Study (DLBS): This study is supported by the Center for Vital Longevity, the
390 University of Texas at Dallas, the University of Texas Southwestern Medical Center, the National
391 Institutes of Health and Aging, AVID Radiopharmaceuticals, the Aging Mind Foundation and the
392 Alzheimer's Association. http://fcon_1000.projects.nitrc.org/indi/retro/dlbs.html
393
394 FIND lab sample. Funded by the Dana Foundation; John Douglas French Alzheimer's
395 Foundation; National Institutes of Health (AT005733, HD059205,HD057610, NS073498,
396 NS058899). http://fcon_1000.projects.nitrc.org/indi/retro/find_stanford.html
397
398 Functional Biomedical Informatics Research Network (FBIRN): Provided by the Biomedical
399 Informatics Research Network under the following support: U24-RR021992, by the National
400 Center for Research Resources at the National Institutes of Health, U.S.A.
401 http://www.birncommunity.org/resources/data/

Lifespan Human Connectome Project in Aging (HCP-Aging): HCP-Aging data were obtained from the National Institute of Mental Health (NIMH) Data Archive (NDA). NDA is a collaborative informatics system created by the National Institutes of Health to provide a national resource to support and accelerate research in mental health. Dataset identifier: http://dx.doi.org/10.15154/1520138. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or of the Submitters submitting original data to NDA. http://nda.nih.gov

International Consortium for Brain Mapping (ICBM). The ICBM (Principal Investigator: John Mazziotta, MD, PhD) was funded was provided by the National Institute of Biomedical Imaging and BioEngineering. ICBM is the result of efforts of co-investigators from UCLA, Montreal Neurologic Institute, University of Texas at San Antonio, and the Institute of Medicine, Juelich/Heinrich Heine University - Germany." https://ida.loni.usc.edu/login.jsp?project=ICBM

Information eXtraction from Images (IXI): Data collected as part of the project EPSRC GR/S21533/02 - http://brain-development.org/ixi-dataset/

F.M. Kirby Research Center neuroimaging reproducibility data (KIRBY-21). Landman, B.A. et al. “Multi-Parametric Neuroimaging Reproducibility: A 3T Resource Study”, NeuroImage. (2010) NIHMS/PMC:252138 doi:10.1016/j.neuroimage.2010.11.047 https://www.nitrc.org/projects/multimodal

Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD): The MIRIAD investigators did not participate in analysis or writing of this report. The MIRIAD dataset is made available through the support of the UK Alzheimer's Society (RF116). The original data collection was funded through an unrestricted educational grant from GlaxoSmithKline (6GKC).
http://miriad.drc.ion.ucl.ac.uk

National Alzheimer's Coordinating Center (NACC): The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P30 AG062428-01 (PI James Leverenz, MD) P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P30 AG062421-01 (PI Bradley Hyman, MD, PhD), P30 AG062422-01 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI Robert Vassar, PhD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P30 AG062429-01(PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P30 AG062715-01 (PI Sanjay

Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD). https://www.alz.washington.edu/

National Database for Autism Research (NDAR): Data were obtained from the National Institute of Mental Health (NIMH) Data Archive (NDA). NDA is a collaborative informatics system created by the National Institutes of Health to provide a national resource to support and accelerate research in mental health. Dataset identifier: http://dx.doi.org/10.15154/1520138. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or of the Submitters submitting original data to NDA. http://nda.nih.gov

Nathan Kline Institute Rockland (NKI-R) sample (NKI-RS) and Enhanced Sample (NKI-RS): Principal support for the enhanced NKI-RS project is provided by the NIMH BRAINS R01MH094639-01. Funding for key personnel also provided in part by the New York State Office of Mental Health and Research Foundation for Mental Hygiene. Funding for the decompression and augmentation of administrative and phenotypic protocols provided by a grant from the Child Mind Institute (1FDN2012-1). Additional personnel support provided by the Center for the Developing Brain at the Child Mind Institute, as well as NIMH R01MH081218, R01MH083246, and R21MH084126. Project support also provided by the NKI Center for Advanced Brain Imaging (CABI), the Brain Research Foundation, the Stavros Niarchos Foundation and the NIH P50 MH086385-S1 (NKI-RS). http://fcon_1000.projects.nitrc.org/indi/pro/nki.html http://fcon_1000.projects.nitrc.org/indi/enhanced/

468 Open access series of imaging studies (OASIS): The OASIS project was funded by grants P50
469 AG05681, P01 AG03991, R01 AG021910, P50 MH071616, U24 RR021382, and R01 MH56584.
470 http://www.oasis-brains.org/
471
472 POWER: This database was supported by NIH R21NS061144 R01NS32979 R01HD057076
473 U54MH091657 K23DC006638 P50 MH71616 P60 DK020579-31 , McDonnell Foundation
474 Collaborative Action Award, NSF IGERT DGE-0548890, Simon's Foundation Autism Research
475 Initiative grant, Burroughs Wellcome Fund, Charles A. Dana Foundation, Brooks Family Fund,
476 Tourette Syndrome Association, Barnes-Jewish Hospital Foundation, McDonnell Center for
477 Systems Neuroscience, Alvin J. Siteman Cancer Center, American Hearing Research Foundation
478 grant, Diabetes Research and Training Center at Washington University grant.
479 http://fcon_1000.projects.nitrc.org/indi/retro/Power2012.html
480
481 Parkinson's Progression Markers Initiative (PPMI): PPMI – a public-private partnership – is
482 funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners,
483 including Abbvie, Allergan, Amathus, Avid Radiopharmaceuticals, Biogen Idec, BioLegend,
484 Bristol-Myers, Celgene, Cenali, Covance, GE Healthcare, Genentech, GlaxoSmithKline, Glolub
485 Capital, Handl Therapeutics, Insitro, Janssen Neuroscience, Eli Lilly and Company, Lundbeck,
486 Merck, Meso Scale Discovery, Neurocrine, Pfizer, Piramal, Prevail, Roche, Sanofi Genzyme,
487 Servier, Takeda, Teva, UCB, Verily, and Voyager Therapeutics. See http://www.ppmi-info.org for
488 further details.
489

Southwest University Adult Lifespan Dataset (SALD): SALD was supported by the National Natural Science Foundation of China (31470981; 31571137; 31500885), National Outstanding young people plan, the Program for the Top Young Talents by Chongqing, the Fundamental Research Funds for the Central Universities (SWU1509383,SWU1509451,SWU1609177), Natural Science Foundation of Chongqing (cstc2015jcyjA10106), Fok Ying Tung Education Foundation (151023) , General Financial Grant from the China Postdoctoral Science Foundation (2015M572423, 2015M580767), Special Funds from the Chongqing Postdoctoral Science Foundation (Xm2015037, Xm2016044), Key research for Humanities and social sciences of Ministry of Education (14JJD880009). http://fcon_1000.projects.nitrc.org/indi/retro/sald.html

University of Wisconsin, Madison (Birn, Prabhakaran, Meyerand) CoRR sample (UWM): Zuo, X.N., et al. (2014). An open science resource for establishing reliability and reproducibility in functional connectomics. *Scientific data, 1*, 140049. doi: 10.1038/sdata.2014.49 http://fcon_1000.projects.nitrc.org/indi/CoRR/html/uwm_1.html

Wayne State EF Dataset: This dataset was supported by National Institute on Aging grants R01-AG011230, R37-AG011230, R03-AG024630 to Naftali Raz, Ph.D. http://fcon_1000.projects.nitrc.org/indi/retro/wayne_EF.html

Yale Low-Resolution Controls Dataset: Scheinost D, Tokoglu F, Shen X, Finn ES, Noble S, Papademetris X, Constable RT. Fluctuations in Global Brain Activity Are Associated With Changes in Whole-Brain Connectivity of Functional Networks. IEEE Trans Biomed Eng. 2016

512 Dec;63(12):2540-2549. Epub 2016 Aug 16.

513 http://fcon_1000.projects.nitrc.org/indi/retro/yale_lowres.html

514 # References

515

516 1. Pfefferbaum A, Rohlfing T, Rosenbloom MJ, Chu W, Colrain IM, Sullivan EV. Variation in
517 longitudinal trajectories of regional brain volumes of healthy men and women (ages 10 to 85
518 years) measured with atlas-based parcellation of MRI. Neuroimage. 2013;65:176-93. doi:
519 10.1016/j.neuroimage.2012.10.008. PubMed PMID: 23063452; PubMed Central PMCID:
520 PMC3516371.

521 2. Crivello F, Tzourio-Mazoyer N, Tzourio C, Mazoyer B. Longitudinal assessment of global
522 and regional rate of grey matter atrophy in 1,172 healthy older adults: modulation by sex and
523 age. PLoS One. 2014;9(12):e114478. doi: 10.1371/journal.pone.0114478. PubMed PMID:
524 25469789; PubMed Central PMCID: PMC4255026.

525 3. Govindarajan KA, Freeman L, Cai C, Rahbar MH, Narayana PA. Effect of intrinsic and
526 extrinsic factors on global and regional cortical thickness. PLoS One. 2014;9(5):e96429. doi:
527 10.1371/journal.pone.0096429. PubMed PMID: 24789100; PubMed Central PMCID:
528 PMCPMC4008620.

529 4. Kruggel F, Turner J, Muftuler LT. Impact of scanner hardware and imaging protocol on
530 image quality and compartment volume precision in the ADNI cohort. Neuroimage.
531 2010;49(3):2123-33. doi: 10.1016/j.neuroimage.2009.11.006. PubMed PMID: 19913626;
532 PubMed Central PMCID: PMC2951115.

533 5. Potvin O, Dieumegarde L, Duchesne S. Normative morphometric data for cerebral
534 cortical areas over the lifetime of the adult human brain. Neuroimage. 2017;156:315-39. doi:
535 10.1016/j.neuroimage.2017.05.019. PubMed PMID: 28512057.

536 6. Potvin O, Dieumegarde L, Duchesne S. Freesurfer cortical normative data for adults
537 using Desikan-Killiany-Tourville and ex vivo protocols. Neuroimage. 2017;156:43-64. doi:
538 10.1016/j.neuroimage.2017.04.035. PubMed PMID: 28479474.
539 7. Potvin O, Mouiha A, Dieumegarde L, Duchesne S. Normative data for subcortical
540 regional volumes over the lifetime of the adult human brain. Neuroimage. 2016;137:9-20. doi:
541 10.1016/j.neuroimage.2016.05.016. PubMed PMID: 27165761.
542 8. Potvin O, Mouiha A, Dieumegarde L, Duchesne S, Alzheimers Disease Neuroimaging I.
543 FreeSurfer subcortical normative data. Data Brief. 2016;9:732-6. doi:
544 10.1016/j.dib.2016.10.001. PubMed PMID: 27830169; PubMed Central PMCID:
545 PMCPMC5094268.
546 9. Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated
547 labeling system for subdividing the human cerebral cortex on MRI scans into gyral based
548 regions of interest. Neuroimage. 2006;31(3):968-80. doi: 10.1016/j.neuroimage.2006.01.021.
549 PubMed PMID: 16530430.
550 10. Klein A, Tourville J. 101 labeled brain images and a consistent human cortical labeling
551 protocol. Front Neurosci. 2012;6:171. doi: 10.3389/fnins.2012.00171. PubMed PMID:
552 23227001; PubMed Central PMCID: PMC3514540.
553 11. Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri
554 and sulci using standard anatomical nomenclature. Neuroimage. 2010;53(1):1-15. doi:
555 10.1016/j.neuroimage.2010.06.010. PubMed PMID: 20547229; PubMed Central PMCID:
556 PMCPMC2937159.

12. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002;33(3):341-55. PubMed PMID: 11832223.

13. Augustinack JC, Magnain C, Reuter M, van der Kouwe AJ, Boas D, Fischl B. MRI parcellation of ex vivo medial temporal lobe. Neuroimage. 2014;93 Pt 2:252-9. doi: 10.1016/j.neuroimage.2013.05.053. PubMed PMID: 23702414; PubMed Central PMCID: PMCPMC3883990.

14. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry. 2014;19(6):659-67. Epub 2013/06/19. doi: 10.1038/mp.2013.78. PubMed PMID: 23774715; PubMed Central PMCID: PMCPMC4162310.

15. Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, et al. Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). Alzheimers Dement. 2005;1(1):55-66. PubMed PMID: 17476317.

16. Ellis KA, Bush AI, Darby D, De Fazio D, Foster J, Hudson P, et al. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. Int Psychogeriatr. 2009;21(4):672-87. doi: 10.1017/S1041610209009405. PubMed PMID: 19470201.

17. Zuo XN, Anderson JS, Bellec P, Birn RM, Biswal BB, Blautzik J, et al. An open science resource for establishing reliability and reproducibility in functional connectomics. Scientific data. 2014;1:140049. doi: 10.1038/sdata.2014.49. PubMed PMID: 25977800; PubMed Central PMCID: PMC4421932.

579 18. Rohr CS, Okon-Singer H, Craddock RC, Villringer A, Margulies DS. Affect and the brain's
580 functional organization: a resting-state connectivity approach. PLoS One. 2013;8(7):e68015.
581 Epub 2013/08/13. doi: 10.1371/journal.pone.0068015. PubMed PMID: 23935850; PubMed
582 Central PMCID: PMCPMC3720669.
583 19. Shafto MA, Tyler LK, Dixon M, Taylor JR, Rowe JB, Cusack R, et al. The Cambridge Centre
584 for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan,
585 multidisciplinary examination of healthy cognitive ageing. BMC Neurol. 2014;14:204. Epub
586 2014/11/22. doi: 10.1186/s12883-014-0204-1. PubMed PMID: 25412575; PubMed Central
587 PMCID: PMCPMC4219118.
588 20. Taylor JR, Williams N, Cusack R, Auer T, Shafto MA, Dixon M, et al. The Cambridge
589 Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI,
590 MEG, and cognitive data from a cross-sectional adult lifespan sample. Neuroimage. 2017;144(Pt
591 B):262-9. Epub 2015/09/17. doi: 10.1016/j.neuroimage.2015.09.018. PubMed PMID: 26375206;
592 PubMed Central PMCID: PMCPMC5182075.
593 21. Mayer AR, Ruhl D, Merideth F, Ling J, Hanlon FM, Bustillo J, et al. Functional imaging of
594 the hemodynamic sensory gating response in schizophrenia. Hum Brain Mapp.
595 2013;34(9):2302-12. Epub 2012/03/31. doi: 10.1002/hbm.22065. PubMed PMID: 22461278;
596 PubMed Central PMCID: PMCPMC4020570.
597 22. Beall EB, Lowe MJ. Isolating physiologic noise sources with independently determined
598 spatial measures. Neuroimage. 2007;37(4):1286-300. doi: 10.1016/j.neuroimage.2007.07.004.
599 PubMed PMID: 17689982.

600 23. Belleville S, LeBlanc AC, Kergoat MJ, Calon F, Gaudreau P, Hebert SS, et al. The
601 Consortium for the early identification of Alzheimer's disease-Quebec (CIMA-Q). Alzheimers
602 Dement (Amst). 2019;11:787-96. Epub 2019/12/04. doi: 10.1016/j.dadm.2019.07.003. PubMed
603 PMID: 31788534; PubMed Central PMCID: PMCPMC6880140.
604 24. Dallas Lifespan Brain Study (DLBS). Available from:
605 http://fcon_1000.projects.nitrc.org/indi/retro/dlbs.html.
606 25. Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD. Decoding subject-driven
607 cognitive states with whole-brain connectivity patterns. Cereb Cortex. 2012;22(1):158-65. Epub
608 2011/05/28. doi: 10.1093/cercor/bhr099. PubMed PMID: 21616982; PubMed Central PMCID:
609 PMCPMC3236795.
610 26. Keator DB, van Erp TGM, Turner JA, Glover GH, Mueller BA, Liu TT, et al. The Function
611 Biomedical Informatics Research Network Data Repository. Neuroimage. 2016;124(Pt B):1074-
612 9. Epub 2015/09/15. doi: 10.1016/j.neuroimage.2015.09.003. PubMed PMID: 26364863;
613 PubMed Central PMCID: PMCPMC4651841.
614 27. Bookheimer SY, Salat DH, Terpstra M, Ances BM, Barch DM, Buckner RL, et al. The
615 Lifespan Human Connectome Project in Aging: An overview. Neuroimage. 2019;185:335-48.
616 Epub 2018/10/18. doi: 10.1016/j.neuroimage.2018.10.009. PubMed PMID: 30332613; PubMed
617 Central PMCID: PMCPMC6649668.
618 28. Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, et al. A probabilistic atlas and
619 reference system for the human brain: International Consortium for Brain Mapping (ICBM).
620 Philos Trans R Soc Lond B Biol Sci. 2001;356(1412):1293-322. PubMed PMID: 11545704.
621 29. IXI Dataset. Available from: http://brain-development.org/ixi-dataset/.

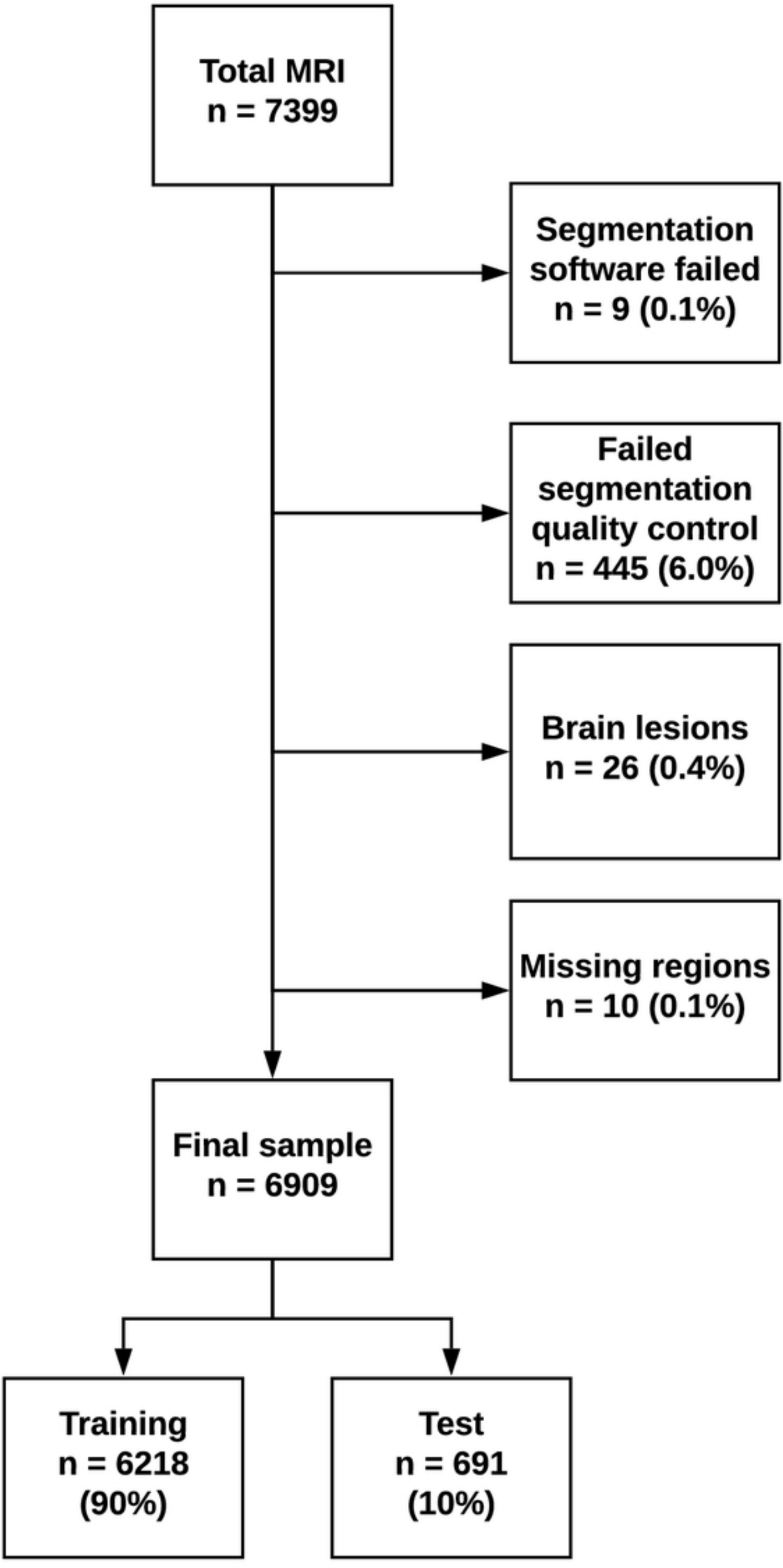
622 30. Landman BA, Huang AJ, Gifford A, Vikram DS, Lim IA, Farrell JA, et al. Multi-parametric
623 neuroimaging reproducibility: a 3-T resource study. Neuroimage. 2011;54(4):2854-66. doi:
624 10.1016/j.neuroimage.2010.11.047. PubMed PMID: 21094686; PubMed Central PMCID:
625 PMC3020263.
626 31. Malone IB, Cash D, Ridgway GR, MacManus DG, Ourselin S, Fox NC, et al. MIRIAD--Public
627 release of a multiple time point Alzheimer's MR imaging dataset. Neuroimage. 2013;70:33-6.
628 Epub 2013/01/01. doi: 10.1016/j.neuroimage.2012.12.044. PubMed PMID: 23274184; PubMed
629 Central PMCID: PMCPMC3809512.
630 32. Beekly DL, Ramos EM, van Belle G, Deitrich W, Clark AD, Jacka ME, et al. The National
631 Alzheimer's Coordinating Center (NACC) Database: an Alzheimer disease database. Alzheimer
632 Dis Assoc Disord. 2004;18(4):270-7. PubMed PMID: 15592144.
633 33. Hall D, Huerta MF, McAuliffe MJ, Farber GK. Sharing heterogeneous data: the national
634 database for autism research. Neuroinformatics. 2012;10(4):331-9. Epub 2012/05/25. doi:
635 10.1007/s12021-012-9151-4. PubMed PMID: 22622767; PubMed Central PMCID:
636 PMCPMC4219200.
637 34. Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, et al. The NKI-
638 Rockland Sample: A Model for Accelerating the Pace of Discovery Science in Psychiatry. Front
639 Neurosci. 2012;6:152. doi: 10.3389/fnins.2012.00152. PubMed PMID: 23087608; PubMed
640 Central PMCID: PMC3472598.
641 35. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. Open Access Series
642 of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented,
643 and Demented Older Adults. J Cogn Neurosci. 2007;19(9):1498-507. PubMed PMID: 17714011.

644 36. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic
645 correlations in functional connectivity MRI networks arise from subject motion. Neuroimage.
646 2012;59(3):2142-54. doi: 10.1016/j.neuroimage.2011.10.018. PubMed PMID: 22019881;
647 PubMed Central PMCID: PMC3254728.
648 37. Parkinson Progression Marker I. The Parkinson Progression Marker Initiative (PPMI).
649 Prog Neurobiol. 2011;95(4):629-35. Epub 2011/09/21. doi: 10.1016/j.pneurobio.2011.09.005.
650 PubMed PMID: 21930184.
651 38. Wei D, Zhuang K, Ai L, Chen Q, Yang W, Liu W, et al. Structural and functional brain scans
652 from the cross-sectional Southwest University adult lifespan dataset. Scientific data.
653 2018;5:180134. Epub 2018/07/18. doi: 10.1038/sdata.2018.134. PubMed PMID: 30015807;
654 PubMed Central PMCID: PMCPMC6049036.
655 39. Moffat SD, Kennedy KM, Rodrigue KM, Raz N. Extrahippocampal contributions to age
656 differences in human spatial navigation. Cereb Cortex. 2007;17(6):1274-82. Epub 2006/07/22.
657 doi: 10.1093/cercor/bhl036. PubMed PMID: 16857855.
658 40. Scheinost D, Tokoglu F, Shen X, Finn ES, Noble S, Papademetris X, et al. Fluctuations in
659 Global Brain Activity Are Associated With Changes in Whole-Brain Connectivity of Functional
660 Networks. IEEE Trans Biomed Eng. 2016;63(12):2540-9. Epub 2016/08/20. doi:
661 10.1109/TBME.2016.2600248. PubMed PMID: 27541328; PubMed Central PMCID:
662 PMCPMC5180443.
663 41. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): Recent evidence and
664 development of a shorter version. Clinical Gerontology: a Guide to Assessment and
665 Intervention. New York: The Haworth Press; 1986. p. 165-73.

666 42. Sherif T, Rioux P, Rousseau ME, Kassis N, Beck N, Adalat R, et al. CBRAIN: a web-based,
667 distributed computing platform for collaborative neuroimaging research. Front Neuroinform.
668 2014;8:54. Epub 2014/06/07. doi: 10.3389/fninf.2014.00054. PubMed PMID: 24904400;
669 PubMed Central PMCID: PMCPMC4033081.
670 43. Augustinack JC, Huber KE, Stevens AA, Roy M, Frosch MP, van der Kouwe AJ, et al.
671 Predicting the location of human perirhinal cortex, Brodmann's area 35, from MRI. Neuroimage.
672 2013;64:32-42. doi: 10.1016/j.neuroimage.2012.08.071. PubMed PMID: 22960087; PubMed
673 Central PMCID: PMCPMC3508349.
674 44. Iglesias JE, Van Leemput K, Bhatt P, Casillas C, Dutt S, Schuff N, et al. Bayesian
675 segmentation of brainstem structures in MRI. Neuroimage. 2015;113:184-95. doi:
676 10.1016/j.neuroimage.2015.02.065. PubMed PMID: 25776214; PubMed Central PMCID:
677 PMCPMC4434226.
678 45. Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, et al. A
679 computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI:
680 Application to adaptive segmentation of in vivo MRI. Neuroimage. 2015;115:117-37. doi:
681 10.1016/j.neuroimage.2015.04.042. PubMed PMID: 25936807; PubMed Central PMCID:
682 PMCPMC4461537.
683 46. Rosas HD, Liu AK, Hersch S, Glessner M, Ferrante RJ, Salat DH, et al. Regional and
684 progressive thinning of the cortical ribbon in Huntington's disease. Neurology. 2002;58(5):695-
685 701. PubMed PMID: 11889230.
686 47. Salat DH, Buckner RL, Snyder AZ, Greve DN, Desikan RS, Busa E, et al. Thinning of the
687 cerebral cortex in aging. Cereb Cortex. 2004;14(7):721-30. PubMed PMID: 15054051.

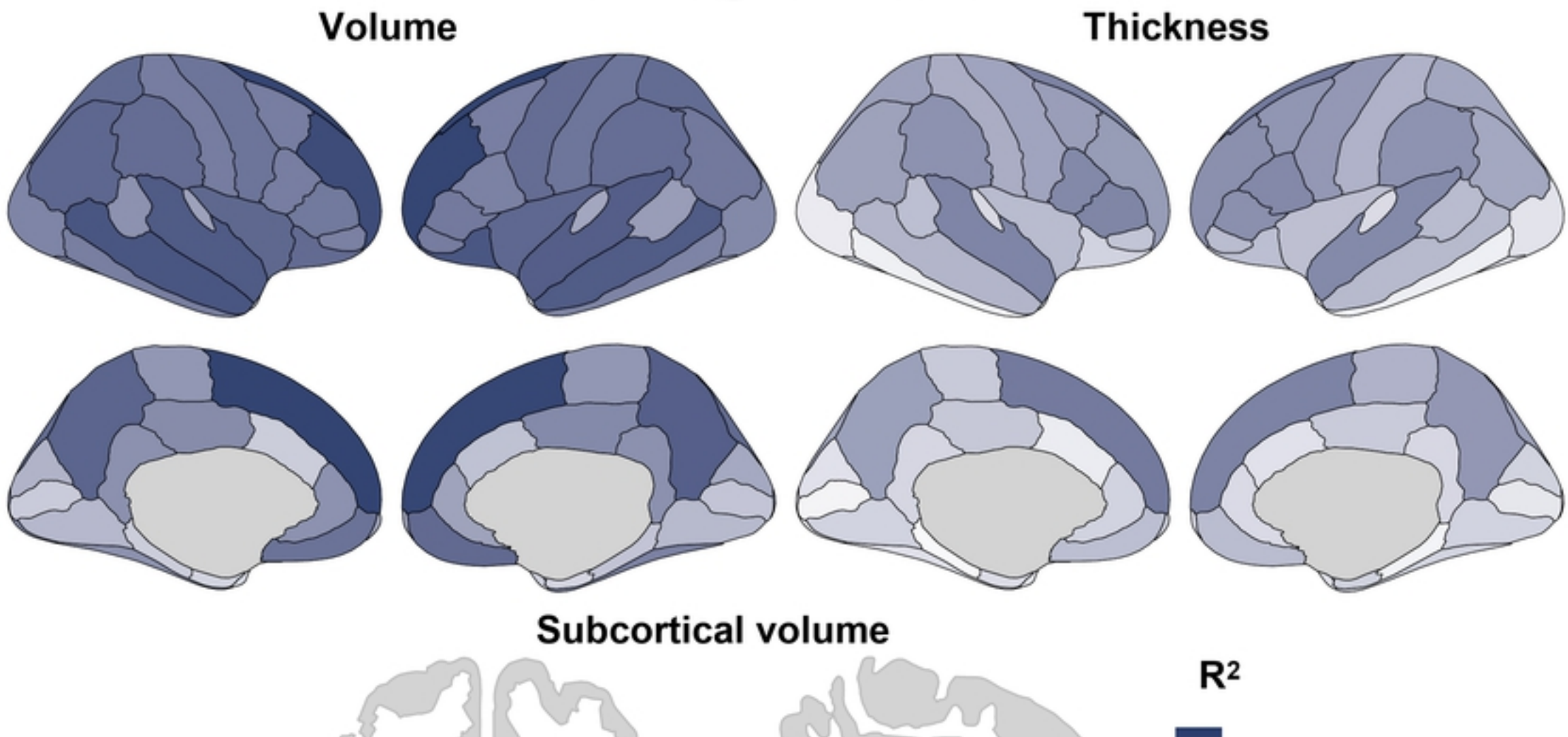
688 48. Kuperberg GR, Broome MR, McGuire PK, David AS, Eddy M, Ozawa F, et al. Regionally
689 localized thinning of the cerebral cortex in schizophrenia. Arch Gen Psychiatry. 2003;60(9):878-
690 88. PubMed PMID: 12963669.
691 49. Buckner RL, Head D, Parker J, Fotenos AF, Marcus D, Morris JC, et al. A unified approach
692 for morphometric and functional data analysis in young, old, and demented adults using
693 automated atlas-based head size normalization: reliability and validation against manual
694 measurement of total intracranial volume. Neuroimage. 2004;23(2):724-38. doi:
695 10.1016/j.neuroimage.2004.06.018. PubMed PMID: 15488422.
696 50. Klapwijk ET, van de Kamp F, van der Meulen M, Peters S, Wierenga LM. Qoala-T: A
697 supervised-learning tool for quality control of FreeSurfer segmented MRI data. Neuroimage.
698 2019;189:116-29. doi: 10.1016/j.neuroimage.2019.01.014. PubMed PMID: 30633965.
699 51. Backhausen LL, Herting MM, Buse J, Roessner V, Smolka MN, Vetter NC. Quality Control
700 of Structural MRI Images Applied Using FreeSurfer-A Hands-On Workflow to Rate Motion
701 Artifacts. Front Neurosci. 2016;10:558. doi: 10.3389/fnins.2016.00558. PubMed PMID:
702 27999528; PubMed Central PMCID: PMCPMC5138230.
703 52. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Data mining,
704 inference, and prediction. Statistics, editor: Springer; 2008. 745 p.
705 53. R Development Core Team. R: A language and environment for statistical computing.
706 Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org2010.
707 54. Grömping U. Relative Importance for Linear Regression in R: The Package relaimpo.
708 Journal of Statistical Software. 2006;17(1):1-27.

709 55. Mowinckel AM, Vidal-Pineiro V. Visualisation of Brain Statistics with R-packages ggseg
710 and ggseg3d. arXiv. 2019. doi: https://arxiv.org/abs/1912.08200.
711 56. Crawford JR, Garthwaite PH, Denham AK, Chelune GJ. Using regression equations built
712 from summary data in the psychological assessment of the individual case: extension to
713 multiple regression. Psychological assessment. 2012;24(4):801-14. doi: 10.1037/a0027699.
714 PubMed PMID: 22449035.
715

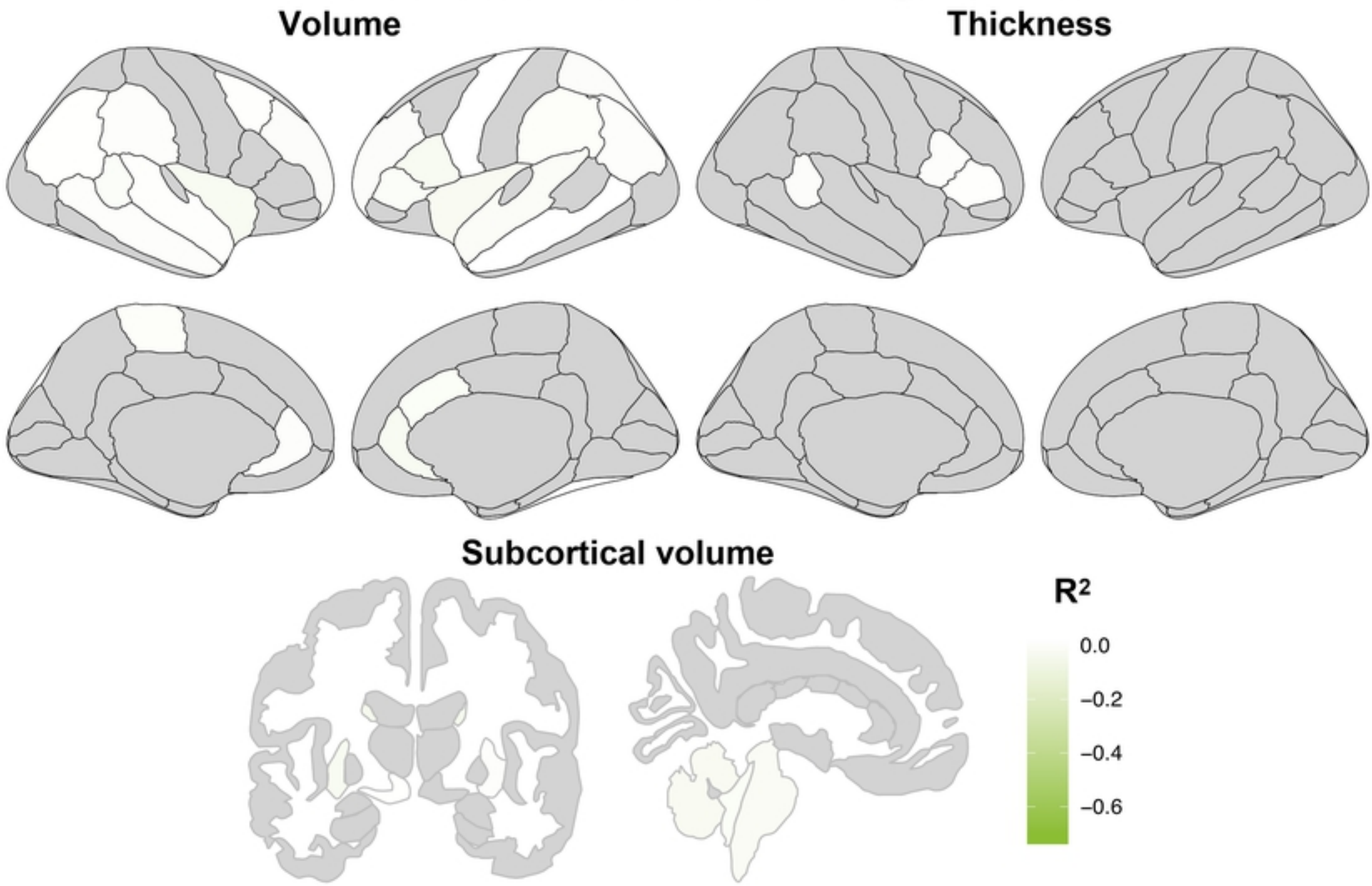

Total MRI
n = 7399
Segmentation software failed
n = 9 (0.1%)
Failed segmentation quality control
n = 445 (6.0%)
Brain lesions
n = 26 (0.4%)
Missing regions
n = 10 (0.1%)
Final sample
n = 6909
Training
n = 6218
(90%)
Test
n = 691
(10%)

Total variance explained

Training 10-fold CV

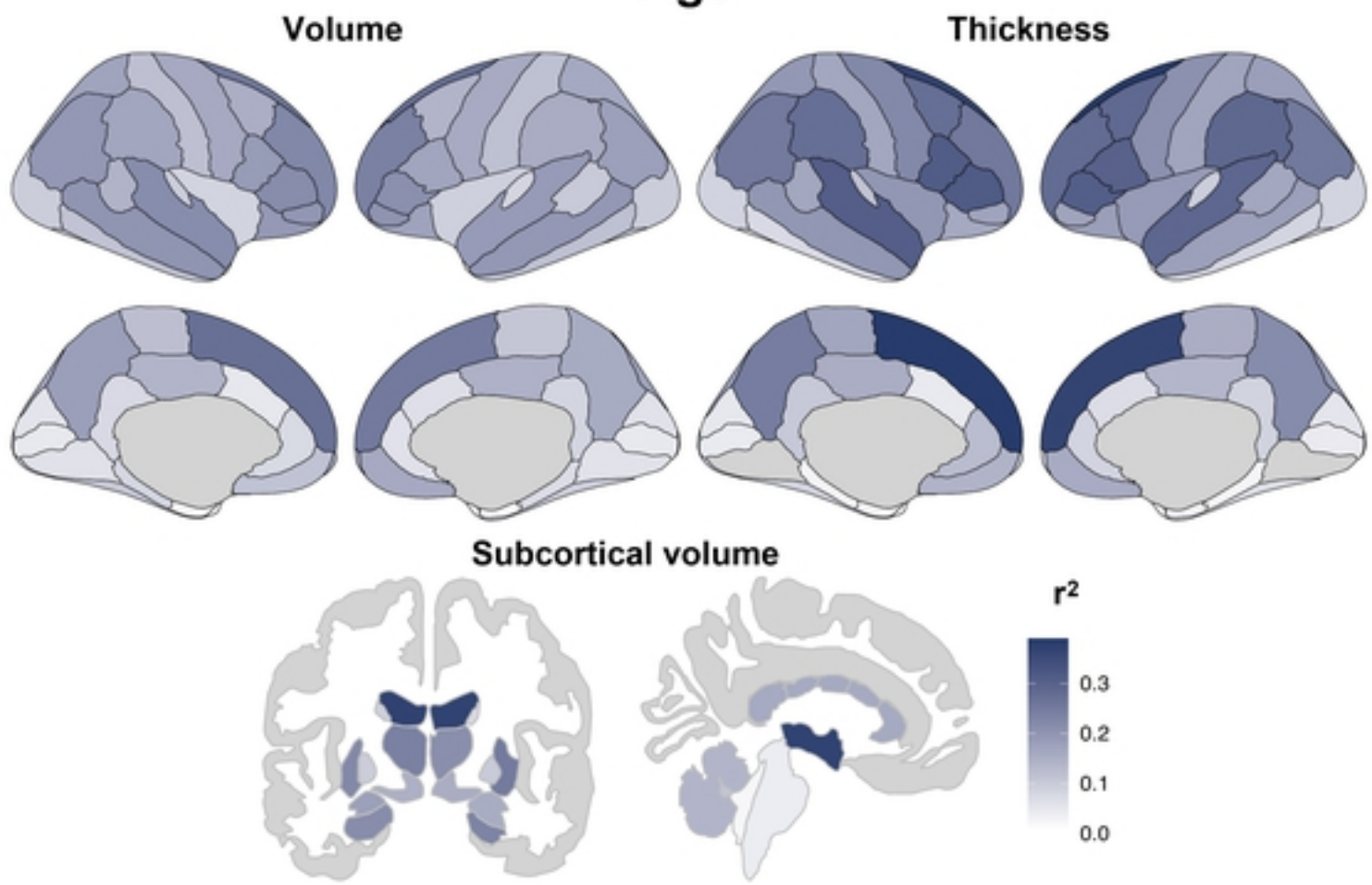


Test validation (test - training values)

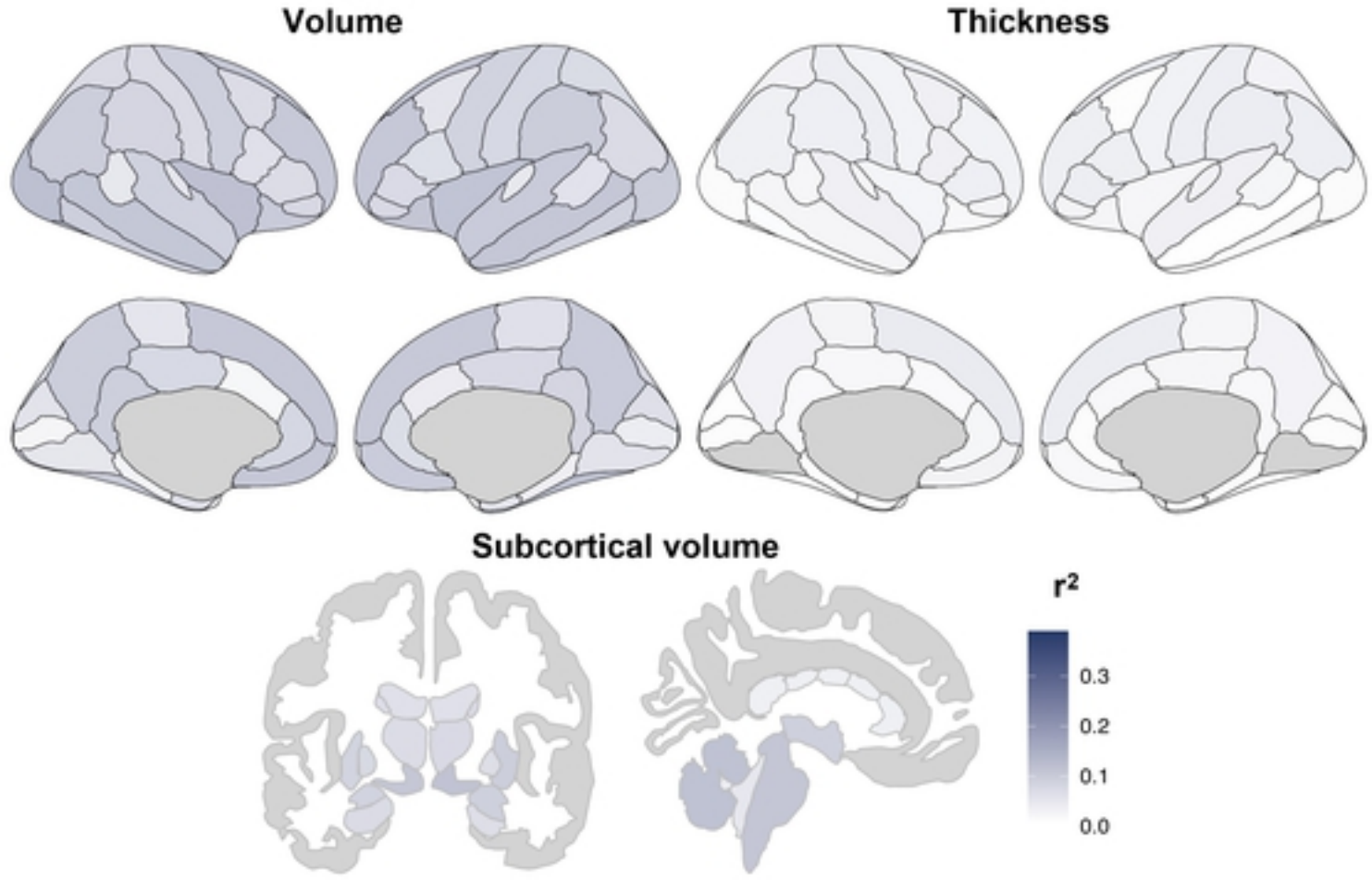


Biological effects

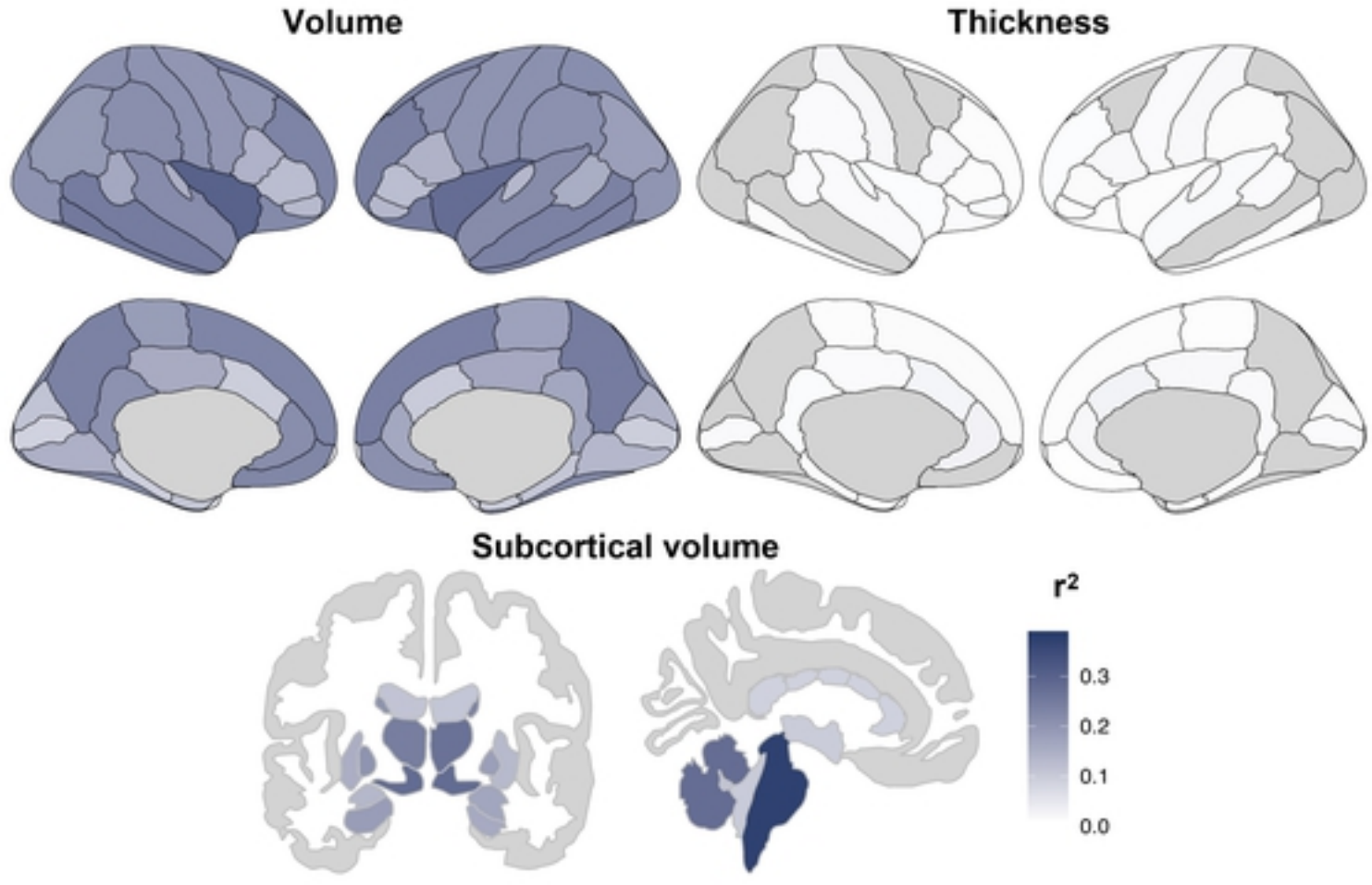
Age



Sex



eTIV



MRI effects

Scanner

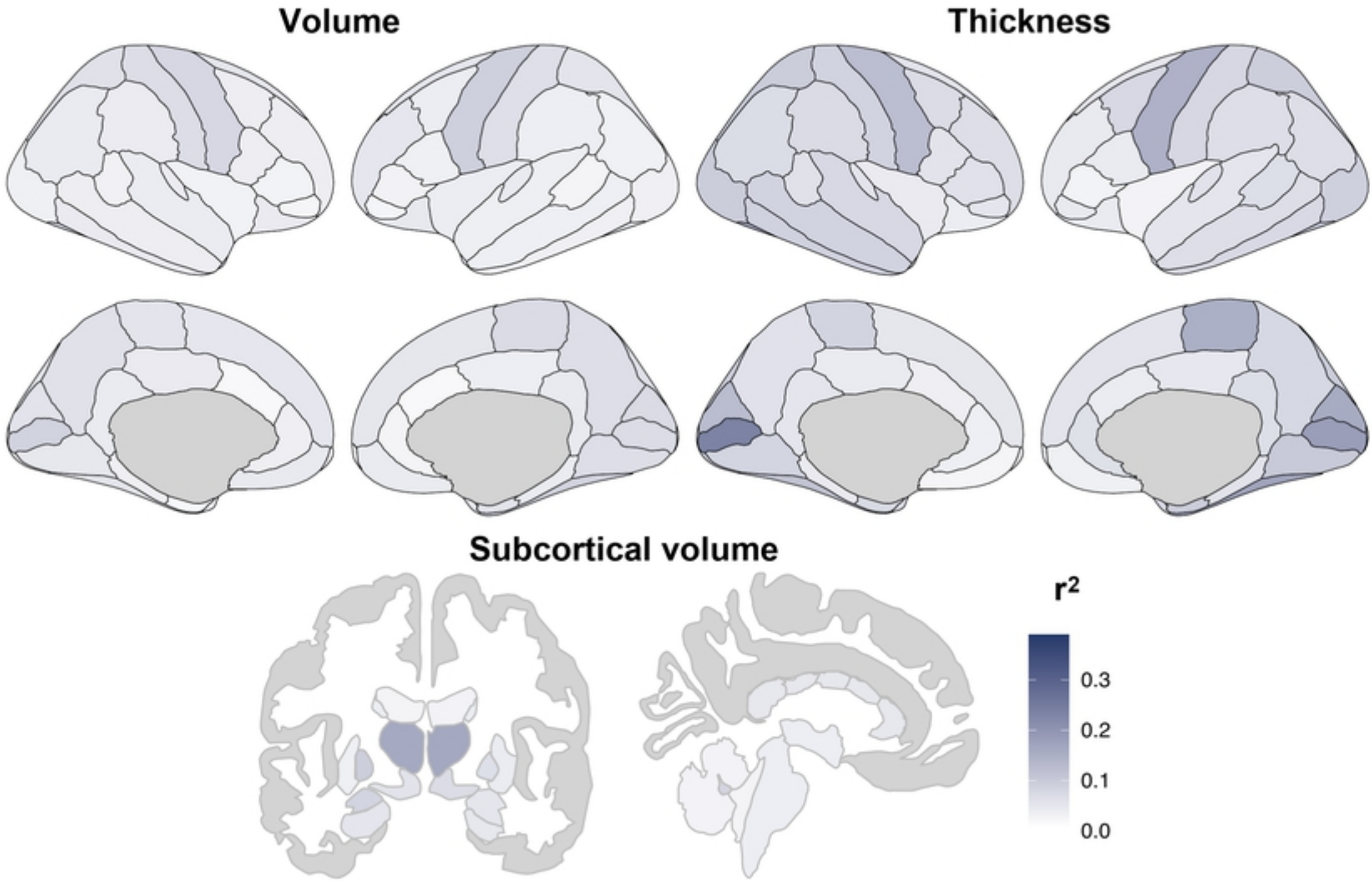
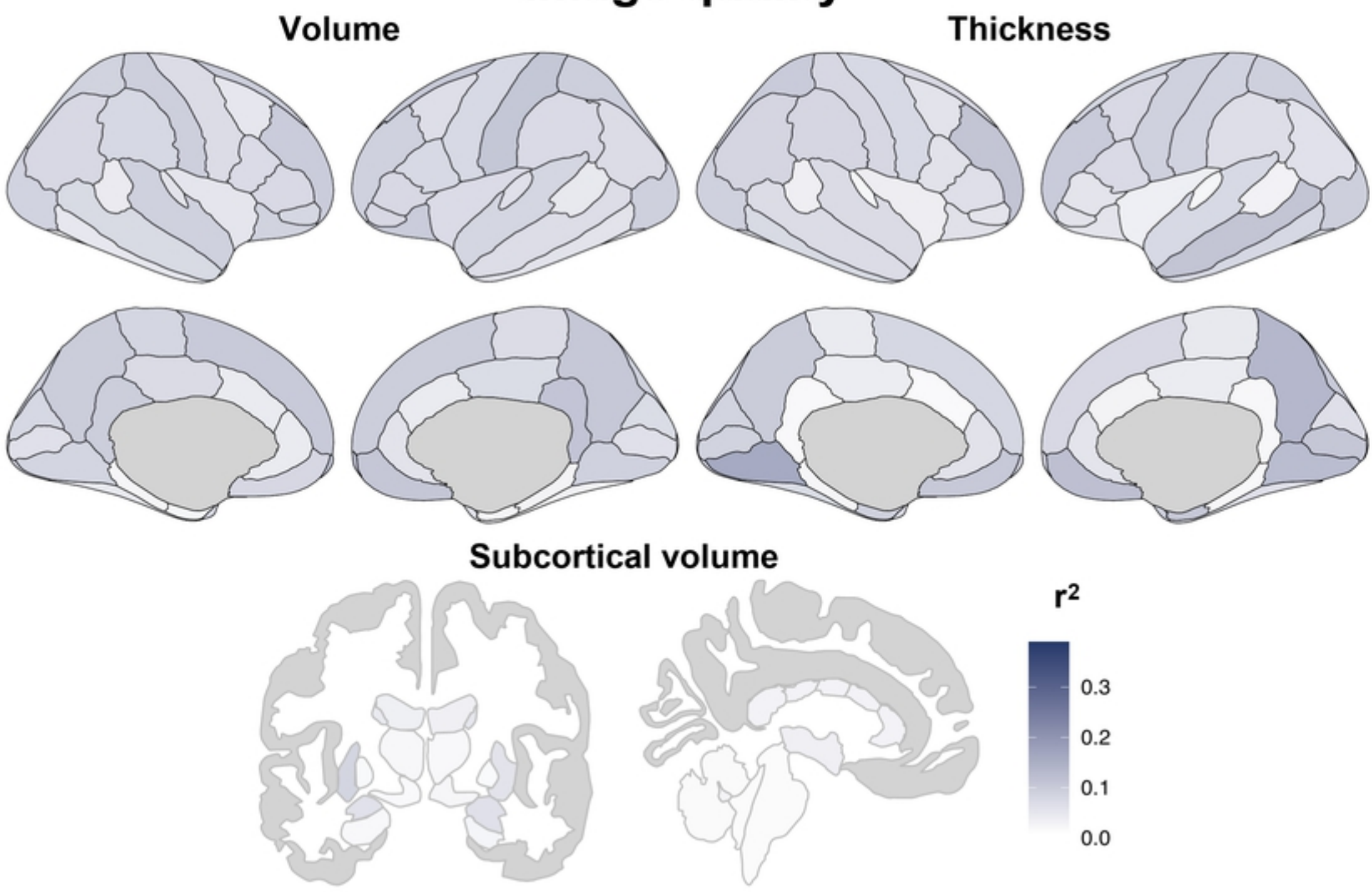
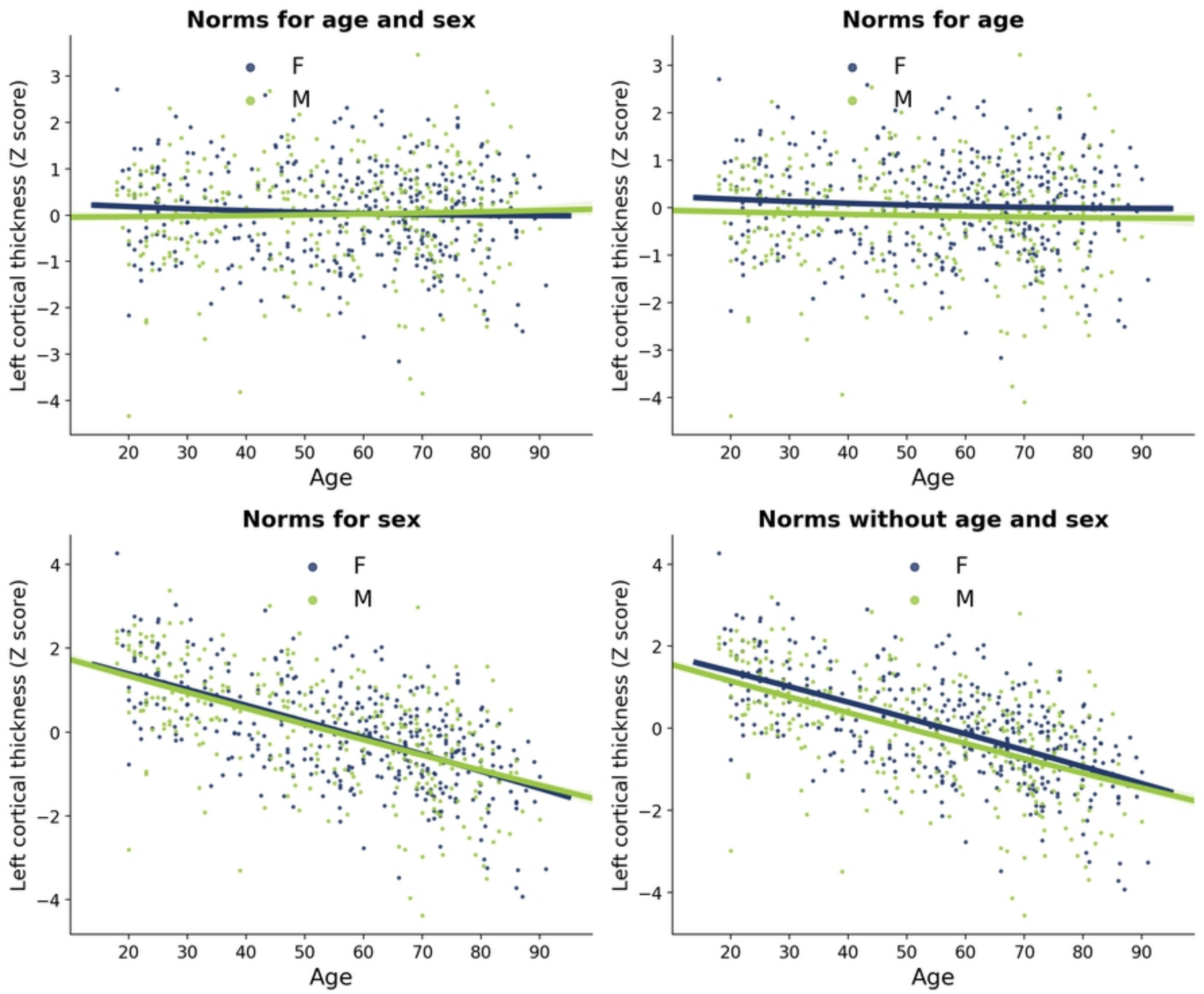


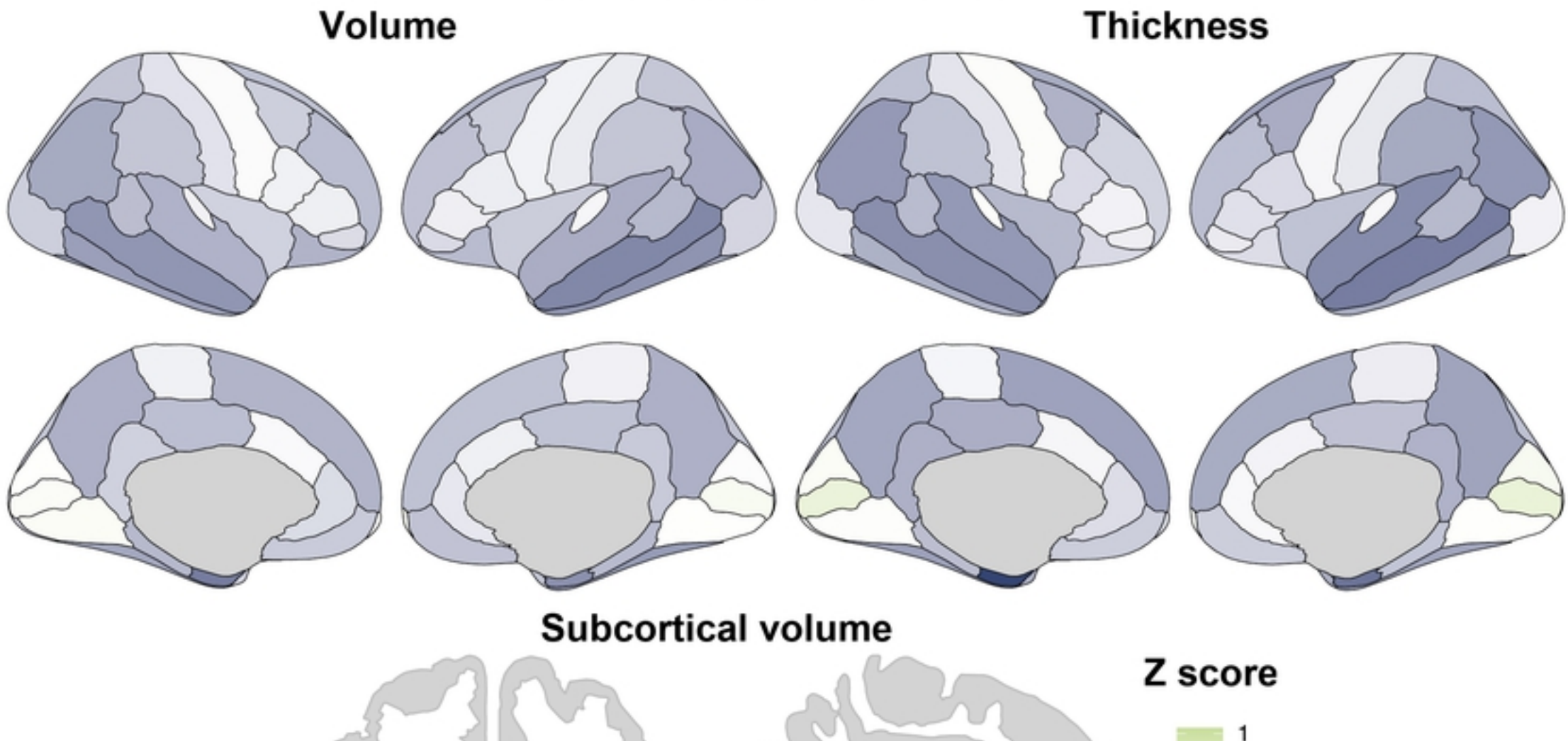
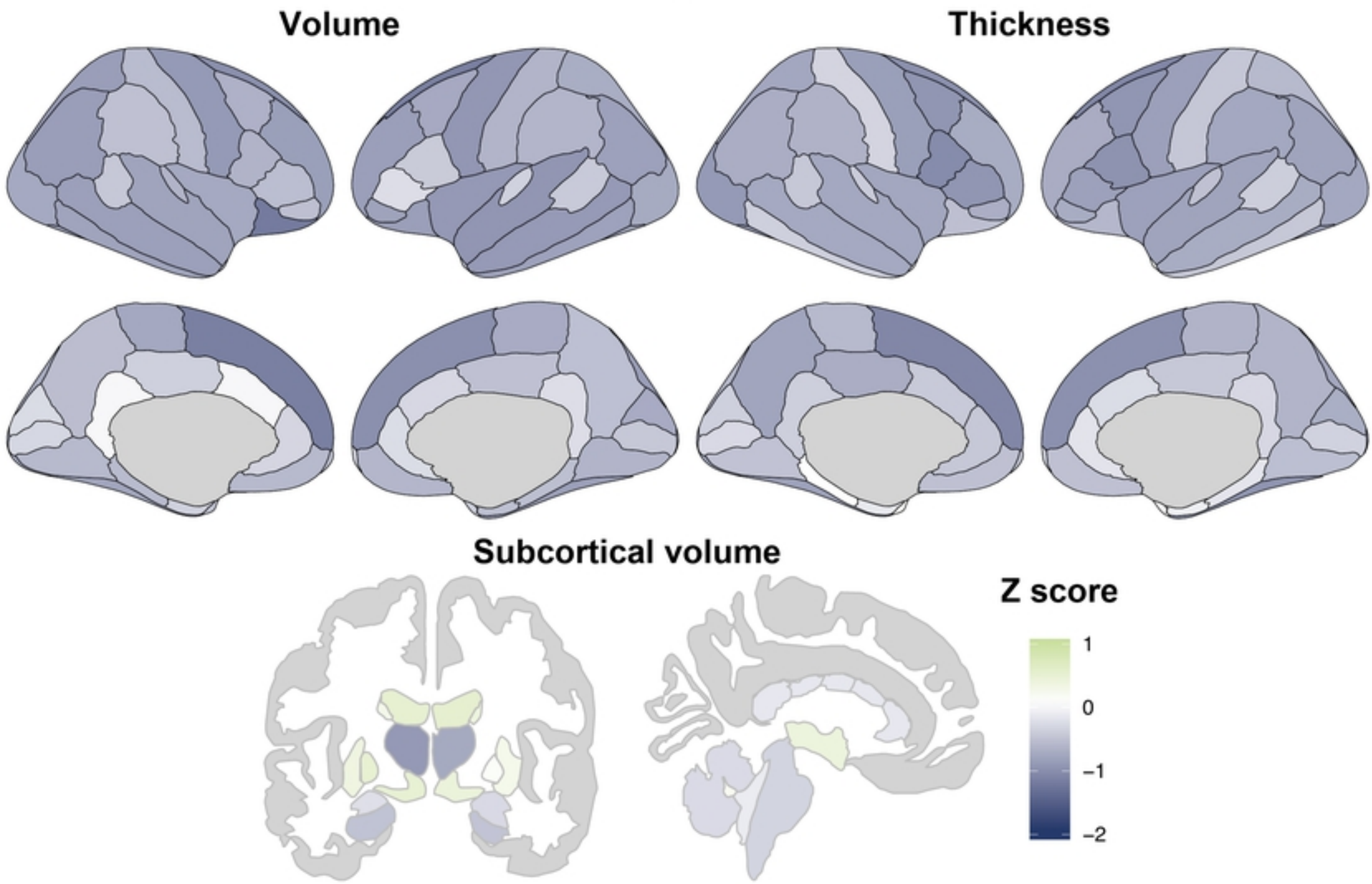
Image quality




Norms for age and sex
Norms for age
Norms for sex
Norms without age and sex
F
M
Left cortical thickness (Z score)
Age

Norms for age and sex

Alzheimer's disease



Schizophrenia