

CIVET-Macaque: an automated pipeline for MRI-based cortical surface generation and cortical thickness in macaques

Claude Lepage^a, Konrad Wagstyl^{a,b}, Benjamin Jung^{c,d}, Jakob Seidlitz^{e,f}, Caleb Sponheim^g, Leslie Ungerleider^c, Xindi Wang^a, Alan C. Evans^a, Adam Messinger^c

^a Montreal Neurological Institute (MNI), McGill University, Montreal, Canada.

^b Wellcome Centre for Human Neuroimaging, University College London, London, UK.

^c Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda MD USA.

^d Department of Neuroscience, Brown University, Providence RI USA.

^e Department of Child and Adolescent Psychiatry and Behavioral Science, Children's Hospital of Philadelphia, Philadelphia PA USA.

^f Department of Psychiatry, University of Pennsylvania, Philadelphia PA USA.

^g Department of Organismal Biology and Anatomy, University of Chicago, Chicago IL USA.

Corresponding author: Adam Messinger, Ph.D.

Email: messinga@mail.nih.gov.

Mailing address: Laboratory of Brain and Cognition, National Institute of Mental Health, 49 Convent Dr., Bldg 49, Room 1B80, Bethesda, MD 20892-4415, USA.

Abstract

The MNI CIVET pipeline for automated extraction of cortical surfaces and evaluation of cortical thickness from *in-vivo* human MRI has been extended for processing macaque brains.

Processing is performed based on the NIMH Macaque Template (NMT), as the reference template, with the anatomical parcellation of the surface following the D99 and CHARM atlases.

The modifications needed to adapt CIVET to the macaque brain are detailed. Results have been obtained using CIVET-macaque to process the anatomical scans of the 31 macaques used to generate the NMT and another 95 macaques from the PRIME-DE initiative. It is anticipated that the open usage of CIVET-macaque will promote collaborative efforts in data collection and processing, sharing, and automated analyses from which the non-human primate brain imaging field will advance.

KEYWORDS: surface-based morphometry, surface registration, rhesus monkey, template, atlas, primate

1. Introduction

The cerebral cortex has a two-dimensional folded topology, tangentially subdivided into cortical areas and radially into layers. Human neuroimaging studies of cortical structure and function have been advanced through the development of automated surface reconstruction techniques (Kim et al. 2005; Bruce Fischl 2012). These techniques have enabled intersubject and intermodal coregistration, compression and analysis of large datasets, and revealed areal and laminar organization in a common surface-based coordinate system.

Cortical meshes enable quantification of morphometry: per-vertex cortical thickness, curvature, and surface area can all be readily measured from mesh representations. The spherical topology of the cortical surfaces allow for intersubject coregistration (B. Fischl et al. 1999; Lerch and Evans 2005; Robinson et al. 2014). Surfaces can be aligned to a common average template, enabling group-comparison studies, and surface-based parcellations such as atlases can be mapped to individual surface meshes for regional analyses, with per-vertex correspondence between the white matter and pial surfaces. Furthermore, coregistering other imaging modalities to the same space enables quantitative sampling of any number of *in vivo* modalities: fMRI, qMRI, PET, EEG, MEG (Weiskopf et al. 2013; Greve et al. 2014; Bonaiuto et al. 2018; Yeo et al. 2011), as well as *post mortem* histological (Wagstyl et al. 2020) and genetic data (Hawrylycz et al. 2012). Once sampled to the mesh, and coregistered to a common template, vertex-based representations allow for intermodal analysis and application of more complex analytical techniques.

Such techniques have been used to advance human neuroimaging, and have the potential to similarly advance NHP neuroimaging. For instance, in humans, surface analysis has been used to parcellate the cortex according to multimodal *in vivo* data from task and resting state fMRI and structural images in the same subjects (Glasser et al. 2016). Radial sampling between gray and white cortical surfaces across the cortex can produce depth-dependent intensity profiles that, with sufficient resolution and contrast, can be used to segment cortical layers (Wagstyl et al. 2020). Surface-based descriptions have also been used to characterize human semantic organization (Huth et al. 2016) and pathology (Lerch et al. 2005). In NHP studies, surfaces can be used to identify similar regions across species and facilitate interspecies comparison (Xu et

al. 2019; Miranda-Dominguez et al. 2014; Donahue et al. 2018). Thus, cortical surface reconstruction techniques provide a powerful framework for neuroimaging.

Several cortical surface reconstructions for macaques have been presented (Koo et al. 2012; Oguz et al. 2015; Donahue et al. 2018; Wang et al. 2018; Autio et al. 2020) and a few new pipelines for macaque cortical surface generation have recently been shared on the PRIME-Resource Exchange (<https://prime-re.github.io/>) (Messinger et al., *this issue*). These efforts have made surface generation less labor intensive. A surface-free approach to evaluate cortical thickness, using the DiReCT volume registration-based method in ANTS (Tustison et al. 2014), has also been extended to the macaque (Calabrese et al. 2015). However, most of the procedures described require both T1 and T2-weighted anatomical scans of sufficient quality, multiple software packages, various manual interventions, and several end-arounds to get human software tools to process non-human primate data. Furthermore, these packages are limited to exemplary scans or small datasets, some without means for surface-based registration for group analyses of large cohorts. The work presented here addresses the need for an open-access dedicated pipeline for non-human primate cortical surface reconstruction, with capabilities comparable to human pipelines.

The CIVET pipeline (Lepage et al. 2017), which was developed at the Montreal Neurological Institute (MNI), is a comprehensive automated pipeline for corticomeric and morphometric analyses of human MR images. The pipeline provides high resolution cortical surface reconstruction from T1-weighted MR images, featuring intersubject surface-based coregistration with per-vertex correspondence between gray and white matter surfaces. In this work, CIVET has been customized for the processing of non-human primate MR images. The present work covers the extension to the macaque monkey and is based on the NIMH Macaque Template (NMT) (Seidlitz et al. 2018).

The availability of a new automated pipeline for processing non-human primate anatomical scans opens new opportunities for analyzing large datasets with their wealth of structural, functional, diffusion and post mortem data. For example, the UNC-Wisconsin Rhesus Macaque Neurodevelopmental Database provides a rich dataset of longitudinal structural and diffusion MRI during early postnatal development in 34 rhesus macaques (Young et al. 2017). For post mortem data, the Allen Brain Institute provides gene expression and histological data

throughout macaque developmental stages, with reference MRI (Bakken et al. 2016). Others have published postmortem diffusion tensor imaging dataset (Calabrese et al. 2015). Recent initiatives such as the PRIMatE Data Exchange (PRIME-DE) (Milham et al. 2018) have helped to harmonize these data, generating large scale neuroimaging datasets that require robust, automated analytical pipelines. Furthermore, adaptation of statistical map repositories like Neurovault (Gorgolewski et al. 2015) have made it easier to share standardized volumetric results in NHPs. Analysis of such data in terms of anatomical areas and morphometry would be enhanced by surface analysis capabilities (Calabrese et al. 2015). Even existing surface-based studies of NHPs may benefit from improved techniques and automation (Koo et al. 2012).

2. Material and Methods

2.1 Subject Information

This study uses anatomical MR scans from 3 sources: the *ex-vivo* scan of the single-subject D99 template (Reveley et al. 2017), 31 macaques from the Central Animal Facility at the National Institute of Mental Health (NIMH, USA), and *in-vivo* T1-weighted (T1w) anatomical scans of 95 macaques obtained from the INDI PRIME-DE consortium (Milham et al. 2018, http://fcon_1000.projects.nitrc.org/indi/indiPRIME.html). All animal procedures were conducted in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals or, in the case of the PRIME-DE scans, with the animal care and use policies of the institution where the data was collected (Table 1).

The rhesus macaques (*Macaca mulatta*) scans from the NIMH, referred to hereafter as the NMT cohort, have been previously described (Seidlitz et al. 2018). The 25 male and 6 female animals were, at the time of the scans, juveniles and adults (5.2 ± 1.9 years) with an average weight of 6.3 ± 1.7 kg. The animals did not have surgery or procedures within the intracranial cavity prior to the scans. However, 14 had a non-metallic post surgically implanted on top of their skull that made the dorsal portions of the skull appear thin or absent in some scans. T1w MR volumes were acquired using the modified driven equilibrium Fourier transform (MDEFT) sequence (Lee et al. 1995; Deichmann, Schwarzbauer, and Turner 2004) on a 4.7T horizontal scanner (Bruker Biospec 47/40). On the day of scanning, animals were placed under food restriction, anesthetized with isoflurane, and placed in the scanner in the sphinx position. Each whole-brain MDEFT scan was acquired over a 40-60 minute session using either a 14 or 16.5 cm diameter

single loop circular head coil (FOV = 96 x 96 x 70 mm³; 0.5 mm isotropic voxels). For most subjects, multiple MDEFT scans were collected consecutively during the session (mean = 2.5 MDEFT scans/subject, range = 1-7). In these cases, subsequent MDEFT scans were linearly (6 parameter rigid-body) registered to the first, and then averaged to create a single volume for each subject.

Site	N (M/F)	Age range (yr)	Scanner Type
1-Newcastle University	12 (10/2)	5-13	Bruker Vertical, 4.7T
2-NIMH (Leopold, Russ)	2 (1/1)	6-7	Bruker Vertical, 4.7T
3-Princeton University	2 (2/0)	3	Siemens Prisma, 3T
4-University of Minnesota	2 (0/2)	10+	Siemens Syngo, 7T
5-UC Davis	19 (0/19)	18-22	Siemens Skyra, 3T
6-NKI	2 (1/1)	6-7	Siemens TIM Trio, 3T
7-Mount Sinai	9 (8/1)	3-8	Philips Achieva, 3T
8-Mount Sinai	5 (5/0)	5-6	Siemens Skyra, 3T
9-Oxford University	20 (20/0)	2-6	n/a, 3T
10-Western Ontario	3 (3/0)	4-7	Siemens Magnetom, 7T
11-Institute of Neuroscience	8 (7/1)	3-5	Siemens TIM Trio, 3T
12-McGill University	1 (0/1)	12	Siemens TIM Trio, 3T
13-East China Normal University-Kwok	4 (4/0)	2-3	Siemens TIM Trio, 3T
14-Aix-Marseille Universite	4 (3/1)	7-8	Siemens Prisma, 3T
15-Netherlands Institute for Neuroscience	2 (2/0)	5	Philips Ingenia, 3T

Table 1: Scanning information for 95 selected T1w scans from the INDI PRIME-DE consortium. Only scans obtained in the sphinx position with complete published information about age and sex were retained.

The scans for the INDI PRIME-DE consortium consisted of *Macaca mulatta* (N=90) and *macaca fascicularis* (N=5) monkeys and were collected at various sites with different scanner types and field strengths. All scans retained for this study (T1w only) were collected in the sphinx position and had published age and sex information (see Table 1). Scans of very poor quality or with obvious lesions were excluded from this experiment.

2.2 Overview

The CIVET pipeline for extraction of cortical surfaces from in-vivo MR images of the human brain has been extended for processing macaque brains, and as such the processing stages for macaques parallel those for the human. Extensions of CIVET to the macaque use the NIMH Macaque Template (NMT) as the standardized space for extracting and characterizing the surfaces of individual monkeys (Seidlitz et al. 2018). The macaque option for CIVET, herein referred to as CIVET-macaque, is invoked by the simple specification of a command line parameter for the choice of the model for the standardized space. CIVET takes as input a T1w scan and registers it to standardized space, herein the NMT (Seidlitz et al. 2018), using an affine transformation. The image is resampled into the NMT space, corrected for non-uniformities, masked, then classified into cerebrospinal fluid (CSF), cortical and subcortical gray matter (GM), and white matter (WM). A white matter surface, per hemisphere, is fitted at the GM-WM boundary using a marching-cubes algorithm, then adjusted to the maximum T1w intensity gradient position, before being expanded to the pial surface. Cortical thickness is calculated as the distance between the two cortical surfaces. Surface-based registration is performed and the cortical thickness maps are blurred and resampled at the common vertex locations of the template for group comparisons and analyses.

The pipeline is based on the *minc* storage format developed at the MNI (Vincent et al. 2016) as the processing unit. The *minc-2* format is based on the HDF5 library with automatic internal file compression for efficient data representation of large volumes (The HDF Group 2016). Cortical surfaces are represented using the *MNI-obj* format

(<http://www.bic.mni.mcgill.ca/Services/HowToWorkWithObjectFiles>). Many tools to manipulate and visualize volumes and surfaces are available in the openly available *minc toolkit* (<http://bic-mni.github.io/#MINC-Toolkit>). An option has also been added to CIVET to export its outputs to the commonly used NIfTI/GIfTI formats, enabling compatibility with a wide range of other analysis tools. The CIVET pipeline is available on the CBRAIN platform (Sherif et al. 2014)

for collaborative studies and sources and binaries for Linux platforms are available via the CIVET project page (https://github.com/aces/CIVET_Full_Project).

Processing Step	Human Processing	Macaque Processing
Stereotaxic template	MNI-ICBM, ADNI, Colin27	NMT
Stereotaxic voxel sampling	0.50mm or 1.0mm	0.25mm or 0.50mm
Volumetric registration (fwhm, simplex, sampling)	length scale 1.0	length scale 0.4
Neck cropping (headheight)	175mm	50mm
Masking (search distance)	length scale 1.0	length scale 0.4
Non-uniformity corrections	spline distance 125 (3T)	spline distance 50 (4.7T)
Classification	priors on human template	priors on NMT
Segmented masks (brainstem, ventricles, cerebellum, etc.)	on human template	on NMT
Surface extraction (fwhm, gradient, dilation/erosion)	length scale 1.0	length scale 0.4
Surface resolution	40,962 or 163,842 vertices	40,962 or 163,842 vertices
Cortical thickness method	tlink, tlaplace, tfs	tlink, tlaplace, tfs
Cortical thickness blurring	30mm fwhm	8mm fwhm
Surface registration (fwhm)	length scale 1.0	length scale 0.4
Surface parcellation	AAL, DKT, lobes	CHARM/D99, lobes

Table 2: Differences in processing steps between the CIVET-human and CIVET-macaque configurations.

While the processing flow in CIVET is the same for humans and macaques, the main implementation issues in the extension of CIVET to macaques are the elaboration of species-specific volumetric and surface templates for registration as well as the scaling of all length-related parameters. Numerous modules of the pipeline have been modified to include a length scale parameter to account for the smaller brain size in the macaque. A linear factor of

0.4 (relative to the human brain) is used for the macaque. Length scales appear in image blurring (full width at half maximum, fwhm), registration (search distances), masking (search radius), surface blurring of cortical thickness (fwhm), morphological operations such as voxel erosion and dilation (distance), non-uniformity corrections (spline distance), and voxel size representation proportional to brain structures. Other algorithms such as tissue classification (histogram-based) and cortical surface extraction (topology based) are unaffected. These changes are summarized in Table 2. Further adaptations of the pipeline to other non-human primate brains should be straightforward by introducing the appropriate length scales and supplying species specific templates.

2.3 Volumetric template

The first step in the extension of CIVET is the selection of a volumetric template for stereotaxic registration and the preparation of its associated atlases. The NIMH Macaque Template (NMT) version 1.2 (asymmetric model), built from the non-linear average of the 31 scans from the NMT cohort, was selected as this registration target (<https://afni.nimh.nih.gov/NMT>; Seidlitz et al. 2018). We note here that, while it is not used in CIVET, the NMT has been recently updated and now includes a symmetric macaque template in Horsley-Clarke stereotaxic coordinates (Horsley and Clarke 1908; Jung et al., *this issue*). The NMT (version 1.2) was created by iterative symmetric diffeomorphic registration methods using the ANTs software package (Avants et al. 2010). First, whole-head single subject images were initially aligned to a standard coordinate space (D99-SL from (Reveley et al. 2017)), and resampled from 0.5 to 0.25 mm isotropic resolution. Then, following a voxelwise N4 bias field correction (Avants et al. 2011) applied on each subject's aligned image, a spatially-unbiased population average was created using the symmetric group-wise normalization (SyGN; (Avants et al. 2010; Love et al. 2016)). The resultant NMT was placed in the neurological orientation, rotated so that, like the D99 atlas, the horizontal plane contained both the superior extent of the anterior commissure and the inferior extent of the posterior commissure (Horsley and Clarke 1908; Saleem and Logothetis 2012), and translated to place its origin (AP 0, SI 0, ML 0) on the midline at the center of the anterior commissure. Intensity values in the NMT were capped to account for the extreme values of the blood vessels, and a final N4 bias field correction was performed.

Atlases on the NMT (version 1.2) are created for its use in CIVET. The template is segmented to label the cerebellum, brainstem, ventricles, amygdala, hippocampus, and subcortical region.

This segmentation is used in CIVET to exclude the brainstem and cerebellum during tissue classification and to mask out the ventricles and subcortical gray areas prior to the extraction of the white matter surface. A brain mask of the template is obtained using a tailored version of FSL *bet* for macaques (Smith 2002). Finally, tissue priors are defined on the NMT (500+ points for each of CSF, GM, and WM). Priors are also defined in the background (BG) outside the skull. These priors are later coregistered to each subject and used for training the tissue classifier. The pre-processing steps of the template are summarized in Figure 1.

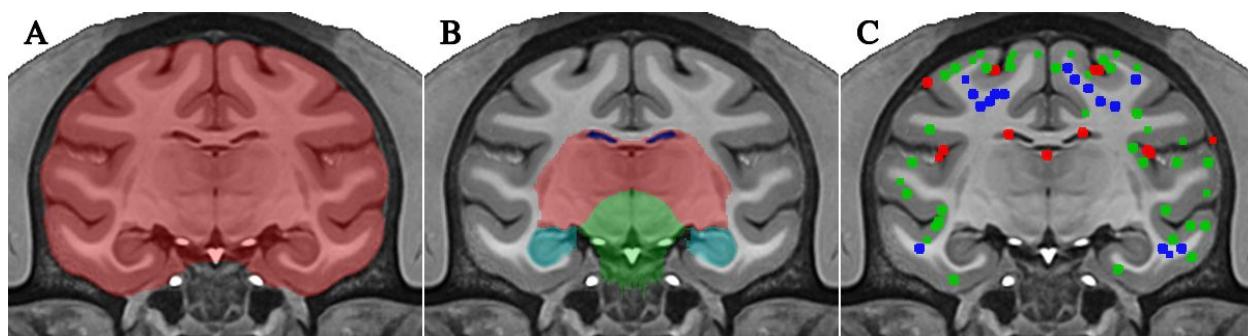


Figure 1: Pre-processing of NMT: A) brain mask (red); B) regional masks for ventricles (dark blue), subcortical region (red), brainstem (green), cerebellum (not visible in this plane), hippocampus (cyan), and amygdala (not visible in this plane); C) subset of the points used to define priors for the three tissue classes: CSF (red), GM (green), WM (blue).

2.4 Registration to standardized template space

Registration to the NMT uses *minctracc* (Collins et al. 1994). Linear registration follows a hierarchical multi-scale approach from coarse to fine image blurring and sampling rates. For the macaque, all length scales in the registration procedure are scaled by 0.4 to account for brain size relative to that of the human. That is, the *fwHM* for image blurring and sampling intervals are scaled proportionally by 0.4. With this scaling, the registration procedure for the human brain achieves optimal performance on the macaque brain. The length scales in the non-linear registration procedure are adjusted in the same way.

Linear and non-linear registration to the macaque template requires the use of a brain mask since head size varies significantly between males and females, unlike humans, resulting in variable scan field-of-views. For instance, the male macaque typically has larger temporalis muscles than the female, resulting in thicker "scalp" (muscle) and head size. In some cases, brains have been pre-processed and de-skulled. Registration, using *minctracc*, proceeds by

specifying the brain mask on the target image only, herein the NMT volume with its associated brain mask, and evaluating the source image, herein the subject's scan, at the sampling points within the target mask, without the need to have a brain mask explicitly defined on the subject. This registration approach has proven to be robust in overcoming the variability in scans due to head size and field-of-view. In the rare case when registration fails, a custom brain mask in native space can be supplied. This custom brain mask can be obtained from automated tools within the CIVET environment or by other external tools. The accuracy of such custom masks should be assessed by the user prior to processing.

The macaque brain is resampled in linear standardized space at the processing voxel resolution of 0.25mm or 0.50mm, which is of comparable resolution relative to the human brain (processing resolution at 0.50mm or 1.0mm).

2.5 Non-uniformity correction

The image, resampled into standardized space, is corrected for non-uniformities using N3 bias field correction (Sled, Zijdenbos, and Evans 1998). The sampling region for N3 is restricted by the brain mask of the volumetric template, as it is for humans, since a reliable custom mask cannot be obtained prior to N3 correction. The parametric spline distance must be entered by the user. This distance varies with the field strength of the scanner and, to some extent, the species' head size. Finally, denoising is performed on the input image to smooth out voxels with local maxima of intensity (100 iterations), thereby avoiding the inclusion of blood vessels in the calculation of the bias field.

2.6 Brain-masking

Brain-masking is based on extensions of FSL *bet* version 1 (Smith 2002). Given the differences in total head size between macaque males and females, it is critical to exclude the scalp prior to calculating the histogram of intensities for establishing the masking thresholds. The most reliable approach was to sample the intensities within the brain mask of the volumetric template, which is permissible in standardized space. Brain extraction a la FSL *bet* thus provides a refinement of the template's brain mask. For some scanning protocols, the blood vessels have very large signal intensities that interfere with convergence and result in an incomplete mask. To prevent this, denoising is performed on the image to smooth out voxels with local maxima of

intensity (100 iterations). Moreover, the brain-masking procedure can incorporate information from complementary MR scans, such as a T2-weighted image, for improving the T1w brain mask. The procedure proved to be very robust at extracting accurate masks on scans of individual subjects (see Figure 2).

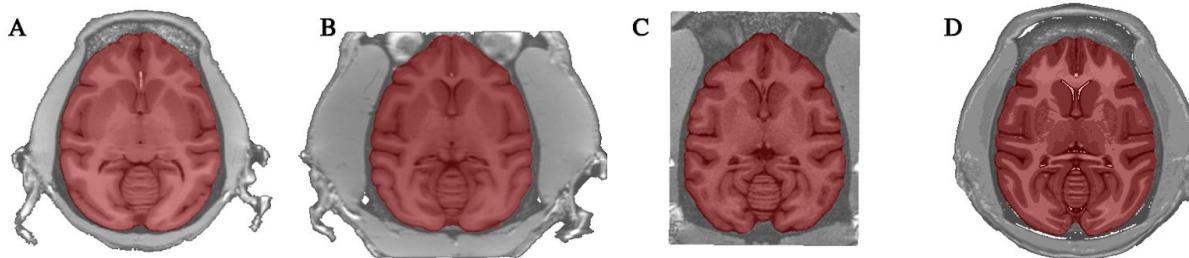


Figure 2: Automated extraction of brain mask in NMT standardized space on sample scans of the NMT cohort, highlighting head size differences in male and female macaques and inter-scan field-of-view variability: A) female macaque; B) male macaque; C) male macaque with cropped field of view; and D) the NMT volume itself (25 males, 6 females). All masks are perfectly extracted regardless of the field of view of the original scan and of the head size of the animal.

2.7 Tissue classification

The tissue classification procedure is identical for both human and macaque scans and relies on the T1w scan, with the possibility of a multi-modal classification should T2-weighted and/or proton density scans be available. The brain is classified into cerebrospinal fluid, gray matter, and white matter using an artificial neural network (ANN) trained on the intensities of the tissue priors (BG, CSF, GM, WM) mapped non-linearly from the NMT volume to the subject (Zijdenbos, Forghani, and Evans 1998). The subcortical region labelled on the NMT allows for the discrimination of cortical from non-cortical gray matter in the specific region. Partial volume estimates are obtained for the tissue classes (Tohka, Zijdenbos, and Evans 2004). Partial volumes for CSF are combined with curvatures (eigenvalues of the Hessian matrix of intensities) in the reconstruction of the CSF skeleton delineating buried sulci which ensures full sulcal penetration of the pial surface during its creation.

2.8 Masking blood vessels

Prior to the extraction of the white matter surface, it is critical to identify T1w hyperintensities, commonly representing major blood vessels and erroneously segmented as white matter by the

classifier, to prevent these spurious white matter voxels from interfering with the process of extracting the white matter surface. The basic assumptions in the detection algorithm are that blood vessels are represented by T1w hyperintensities with a line morphology propagating in CSF within sulci. Sulcal lines are identified using a watershed algorithm in non-white voxels, then dilated partly through gray matter and, consequently, through blood vessels, but not white matter. The trace of the Hessian matrix of second derivatives (i.e., the sum of the Hessian's three eigenvalues) is used as an approximation of the local curvatures to detect one-dimensional line features. Thresholds for the expected signal intensity in white matter and for the trace of the Hessian are used to discriminate lines structures in sulci from blood vessels. Figure 3 illustrates blood vessels detected on a single scan.

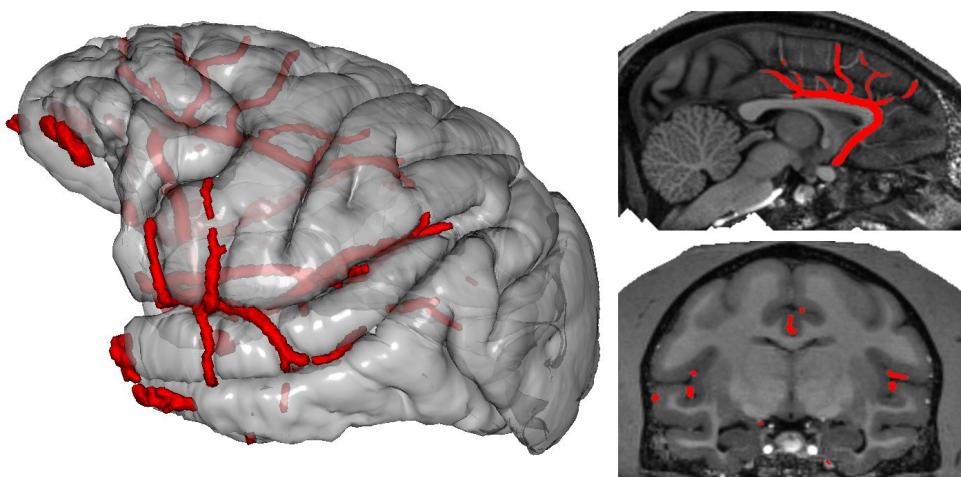


Figure 3: Major blood vessels identified from T1w hyperintensities for one NMT subject. Left: representation of blood vessels (red) on a semi-transparent lateral view of the pial surface. **Right:** T1w hyperintensities revealing the blood vessels in a sagittal and coronal slice through the volume. The vessel-like hyperintensities are discounted during extraction of the white matter surface.

2.9 Cortical surface extraction

A white matter mask is created from the tissue classification, masked for the brain. Using the segmented regional masks on the NMT volume non-linearly transformed to the subject, the ventricles and subcortical gray matter are filled in as white matter, the hippocampus and the amygdala are masked out, and the brainstem and cerebellum are excluded for the purpose of extracting the white matter surface. The brain, in NMT space, is then split into the left and right hemispheres, at the mid-plane wall, and the white matter surfaces are extracted using a marching-cubes algorithm (Lepage et al. 2017). Spherical topology of the cortical mantle is

preserved by advancing the surface, in voxel space, outside-in from a sphere enclosing the hemisphere, hence the requirement for masking off the blood vessels obstructing the progression of the surface. Mesh adaptation operations (vertex smoothing, edge collapsing and swapping) are used to efficiently reduce the surface mesh to a manageable size (around 140,000 vertices) for further processing. The decimated white matter surface is inflated to a sphere and resampled at a fixed number of vertices on a canonical sphere obtained by refinement of the icosahedron (40,962 vertices). The resampled surface is then registered to the average white matter surface template (see subsection below) for mapping the surface region on which the intensity gradient correction is disabled, either because it is meaningless (cut through the corpus callosum and brainstem on the medial wall) or misleading (e.g., cavity of the hippocampus and amygdala, insula region with low gray-white contrast in proximity of the claustrum). The white matter surface is then adjusted to the maximum intensity gradient (magnitude of first derivative of T1w intensity normal to the surface).

The pial surface is obtained by expanding the white matter surface radially outwards up to the pial boundary using the CLASP algorithm (Kim et al. 2005). Consequently, a per-vertex correspondence exists linking the vertices of the white matter and pial surfaces, from which the mid-surface is obtained by averaging the position of the linked vertices. Finally, during the construction process, the right surface is obtained by mirroring the canonical sphere before resampling, thus establishing a direct correspondence between vertices across hemispheres at mirror-symmetric positions.

No special adjustments were made for macaque surface extraction except for scaling (by 0.4) the search distance for the gradient correction and the *fw hm* for blurring volumes. The surfaces can also be extracted at high-resolution (163,842 vertices).

2.10 Surface-based registration

Surface-based registration is used to find the correspondence between the cortical surface of an individual subject to a surface template (Robbins 2003). The registration operates on a spherical representation of the surfaces and the mesh alignment is driven by a *depth potential* term, akin to sulcal depth (Boucher, Whitesides, and Evans 2009). Thus sulci are aligned to sulci and gyri are aligned to gyri. Following registration, any map defined on the subject's surface can be

interpolated at the sampling vertices of the reference surface template. Resampling of surface maps, for instance cortical thickness, enables vertex-wise group comparisons across subjects of a population. Furthermore, regional analyses can be performed based on parcellations defined on the template. The generalization from human to macaque required accounting for head size by scaling the fwhm for blurring the sulcal depth term on the surface.

Surface-based registration uses a white matter surface template for the gradient correction adjustment of the white matter surface and a mid-surface template for resampling surface maps such as cortical thickness, curvature, etc. The next subsections describe the creation of these average white and mid-surface templates as the targets for surface-based registration and the application of the D99 and CHARM parcellation on these surface templates.

2.11 Average surface template

The average surface template is generated from the NMT dataset (N=29, 2 subjects were rejected -- one with a lesion, the second one due to abnormal ventricles) as the target for surface-based registration. Like for the human brain, the average surface is constructed by averaging the surfaces of the individual subjects, as opposed to using the cortical surfaces extracted on the average volume of all subjects. But, unlike the human brain, the average surface is generated in the non-linear standardized space, since subject-to-subject sulcal variability is low for the macaque. That is, the non-linear transformation from the subject's volume to the NMT template is applied to the subject's surfaces before averaging them, thus reducing geometrical variance. This low-frequency non-linear transformation essentially corrects for gross distortion, like global twist, while retaining micro shape differences describing biological variability (see Figure 4). Non-linear warping was also applied in the construction of the NMT volume (Seidlitz et al. 2018).

The creation of the average surface is cyclic: CIVET requires the existence of an average surface as a template for surface-based registration, but surfaces must exist prior to averaging them. An iterative process is followed by first averaging unregistered surfaces and then using this average as the target for registration on the next iteration, and so on. The initial average is made possible since all surfaces have the same number of vertices, via sphere-to-sphere interpolation of the inflated surface to a canonical sphere, and that coordinates are aligned in stereotaxic space (Lyttelton et al. 2007).

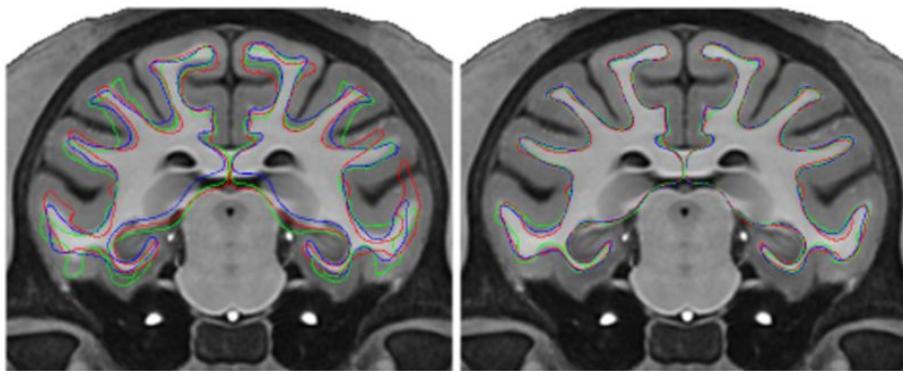


Figure 4: View of white matter surfaces for 3 NMT subjects in linear (affine) space on the left and in non-linear space on the right. Averaging in non-linear space eliminates shape differences while retaining the full cortical morphology of the macaque brain during the creation of the average white matter surface. The surfaces are overlaid onto the T1w NMT. The colored lines correspond to the three subjects.

The generation of the average white matter surface using the marching-cubes algorithm is computationally intensive since the algorithm requires repeated processing steps at each averaging iteration, namely, extraction, mesh decimation, inflation to sphere, resampling, and surface-based registration in order to apply the mask for the gradient correction. This mask must be defined on an intermediate average surface used as the registration target, which itself evolves as the average surface is iteratively converged. As such, the convergence of the individual white matter surfaces depends on the average surface at each iteration and these white matter surfaces must therefore be recomputed in full at each iteration, making the overall process lengthy. To accelerate the convergence process, the intensity gradient correction is disabled for the first 5 iterations of the creation of the average white matter surface. The gradient mask is defined on the intermediate average white matter surface and frozen through 5 more global iterations of surface averaging, with the intensity gradient correction enabled. Full convergence is rapidly achieved (see Figure 5).

The creation of the average mid-surface is much simplified since the extraction of the cortical surfaces does not need to be repeated at each iteration. On start-up, the mid-surfaces are initially registered to the average white matter surface to align the vertices between the two surfaces. Ten global iterations are used in the process, but convergence is achieved after only a few iterations (see Figure 5).

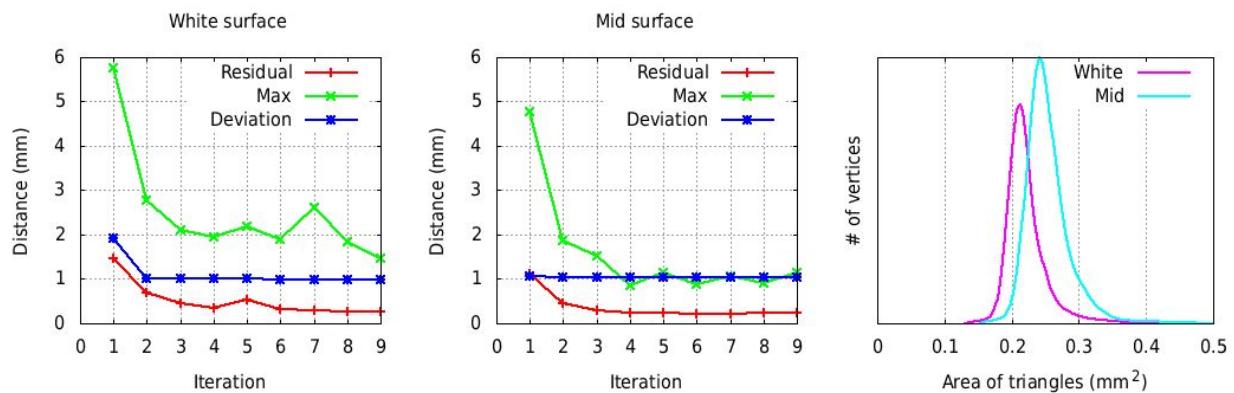


Figure 5: Convergence of the average symmetric white matter (left panel) and mid-cortical (middle panel) surfaces for the NMT cohort ($N=29$). The residual represents the vertex-wise mean displacement in the average surface from one iteration to the next and should tend to zero as the vertices are no longer moving. Max represents the largest vertex-wise displacement from one iteration to the next. Deviation represents the mean vertex-wise displacement for the $N=29$ subjects (58 surfaces) and is a measurement of inter-subject shape variability. The rightmost panel shows the histogram of the area of the triangles on the final average white matter and mid-cortical surfaces, indicating a nearly uniform point sampling over the surface.

For the symmetric averages, the right surfaces are mirrored as left surfaces before being averaged. For both white matter and mid average surfaces, the vertices are equidistributed on the surface to obtain a surface with nearly equal-size triangles, with the goal to have uniform sampling over the entire cortical sheet (see Figure 5). The area equalization process is performed at the end of each averaging iteration and consists of moving each vertex to the centroid of its neighboring vertices, weighted by the area of their connected triangles.

The average white matter surface is shown in Figure 6. The variability in shape (mean vertex displacement) across the $N=29$ subjects is highest in the parieto-occipital area and the medial and ventral occipital lobe. The average mid-cortical surface exhibits the same properties as the average white matter surface.

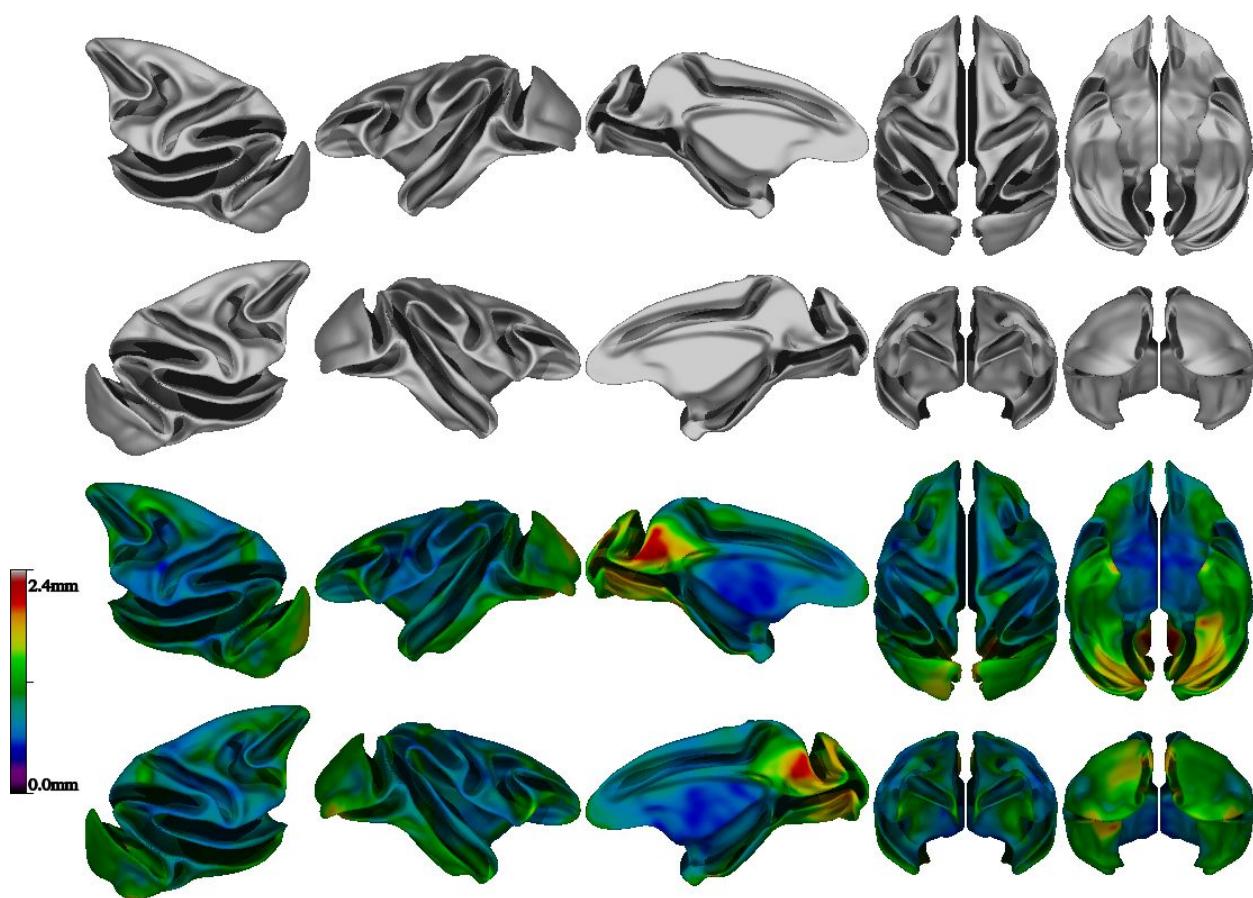


Figure 6: View of the average white symmetric surface colored by mean vertex displacement for N=29 subjects of the NMT dataset. Higher variability in shape was observed in the parieto-occipital area (visual areas V6/V6A) as well as in the occipital lobes, areas of substantial inter-subject differences in macaques.

2.12 Cortical morphometry

The cortical surfaces, extracted in standardized space, are transformed back to native space and cortical thickness is evaluated as the distance between the white matter and pial surfaces.

Three cortical thickness metrics are available in CIVET:

1. *tlink*: distance between linked vertices between the white matter and pial surfaces (sensitive to mesh distortion during expansion of the pial surface, which gives rise to an overestimation of cortical thickness);
2. *tlaplace*: distance along radial lines defined by a Laplacian field between the white matter and pial surfaces (not sensitive to mesh distortion) (Lerch and Evans 2005);
3. *tfs*: symmetric average of projected distance from white matter-to-pial and pial-to-white matter surfaces (like FreeSurfer).

Other metrics are also produced such as vertex-wise surface areas and volumes, and mean and Gaussian curvatures. These measurements are blurred isotropically on the mid-cortical surface in native space (Boucher, Whitesides, and Evans 2009). Default values for surface blurring (*fw hm*) are proportionally less for the macaque than for the human (e.g. *fw hm* of 8mm). All measurements are resampled on the average NMT symmetric surface template as a basis for group comparison analyses.

2.13 CHARM surface atlas

The D99 subject was processed with CIVET-macaque to obtain the D99 surface parcellation (Reveley et al. 2017). This post-mortem brain scan has no skull, a T1w-like contrast, and has been pre-processed: non-uniformities have been corrected and the scan has been symmetrized and transformed to a standardized space. A value of 0 was used as the spline parametric distance to disable N3 and the blood-vessel algorithm has been disabled as well (because the scan was acquired *ex vivo*). Surface-based registration was used to register the D99 mid-surfaces to the NMT symmetric surface template to ensure a direct vertex-to-vertex correspondence. The resampled D99 mid-cortical surfaces were then intersected with the D99 volumetric segmentation (Reveley et al. 2017) to obtain parcellation labels, which mapped directly to the NMT symmetric surface template (see Figure 7). A few minor morphological operations on the extracted labels were necessary to enforce symmetry and to correct for mis-labeled and unassigned vertices arising when intersecting the surfaces with the labelled volume. Using the final D99 labels, each vertex was assigned 6 additional anatomical labels; one for each level of the Cortical Hierarchical Atlas of the Rhesus Macaque (CHARM) (Jung et al., *this issue*). The 6 anatomical levels of CHARM range from a lobar parcellation at the coarsest level (4 ROIs) to nearly the full set of D99 cortical areas at the finest level (139 ROIs). Figure 8 illustrates the 6 parcellation levels of CHARM. The CHARM parcellation can be used for regional analyses of cortical morphometry at multiple spatial scales.

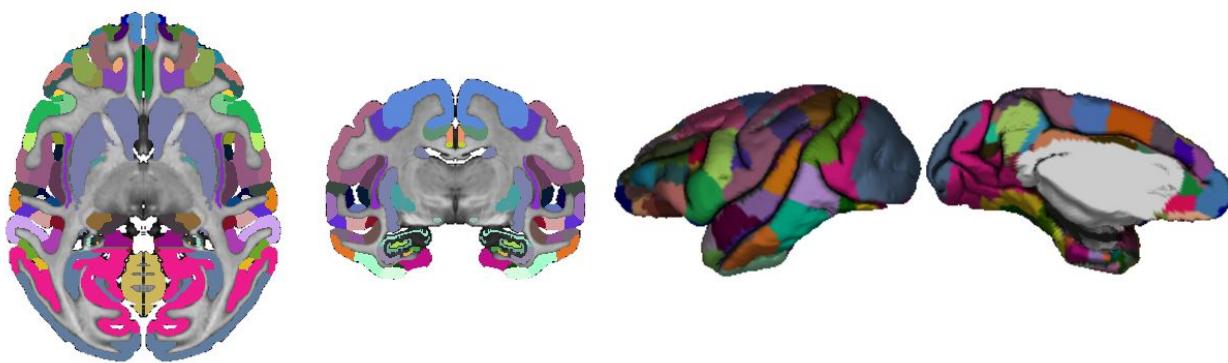


Figure 7: D99 volumetric segmentation defined on the anatomical scan of the D99 subject, transposed to the D99 registered pial surface.

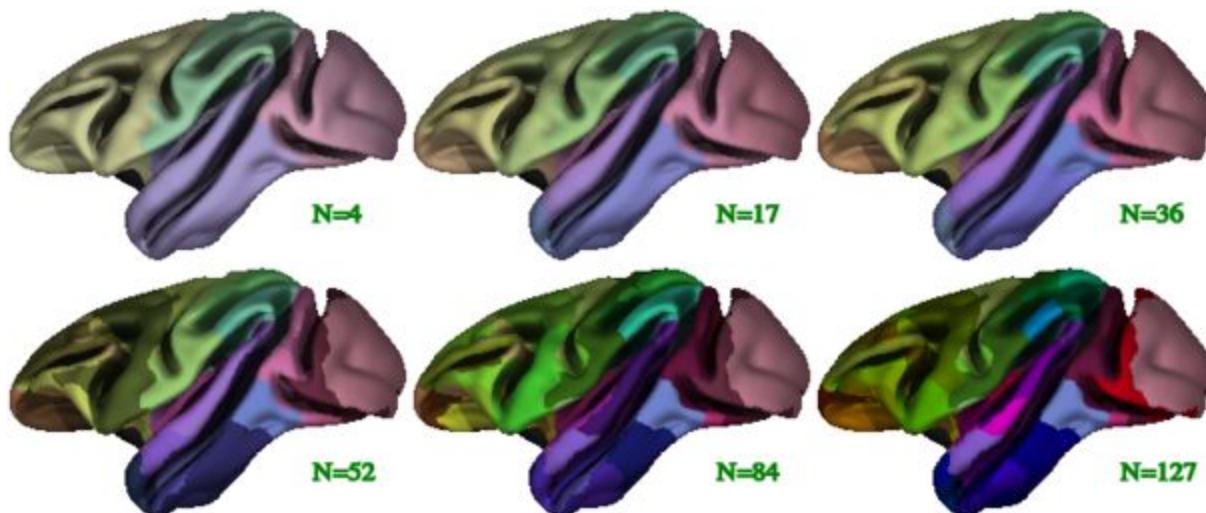


Figure 8: Illustration of the 6 levels of the symmetric CHARM parcellation viewed on the average symmetric mid-cortical surface for the NMT dataset. The CHARM parcellation is used in regional analyses by means of surface-based registration of individual subjects to the NMT surface template.

2.14 Quality control

CIVET produces a series of images to monitor the main processing stages of each subject. These images provide an efficient means for visual inspection of volumetric registration, tissue classification, surface extraction and convergence, surface-based registration, and cortical thickness. Moreover, several metrics are extracted (e.g. brain volume, mean cortical thickness, tissue percentages, etc.) and tabulated into a global csv file for detection of outliers within the population. Failed runs can be systematically identified at this level of quality control.

3. Results

3.1 CIVET preprocessing

The reliability of CIVET-macaque pipeline's volumetric registration and brain-masking was tested on the T1w scans of the 31 macaques in the NMT cohort (Seidlitz et al. 2018) and 95 additional T1w scans of macaques obtained from the INDI PRIME-DE consortium (Milham et al. 2018). CIVET was run with the *without-surface* option, terminating after the stereotaxic registration, brain masking, and tissue classification stages. A value of 50 was used as the spline distance for N3 for all PRIME-DE scans, regardless of the field strength of the scanner, and a value of 75 for the NMT scans. A head height of 50mm, measured from the top of the head downwards, was used to crop the image to remove the neck for the purpose of registration.

Upon visual quality control, all brains were accurately registered to the NMT, without exception. The brain masking stage was successful for all subjects, except those from Site 10 for which the mask leaked into the scalp. For those 3 scans, the bone marrow inside the skull revealed signal intensities similar to GM and WM and was fused with cortex, without a visible inner bone border, causing the mask leakage. (The use of the T2w scan to correct the mask was preempted since the temporal lobes were truncated by the scanner field of view.) In a few other cases, from other sites, the mask touched the edge of the orbits or the bottom half of the cerebellum was truncated from the field of view of the native scan, but such small defects do not have any incidence on cortical surface extraction. Figure 9 reports the volume of the brain masks in standardized space. Agreement of the individual brain mask volumes with that of the NMT (red line) confirms successful registration to the NMT and brain extraction. The few outliers from site 10 are clearly visible in Figure 9. The mean stereotaxic brain mask volume for the PRIME-DE subjects, excluding the 3 outliers from site 10, is $92.9 \pm 3.0\text{cc}$, in agreement with the volume of the NMT brain mask of 92.8cc. Figure 10 shows the scan acquisition voxel size within and across sites. Scans for which all voxel dimensions were large (sites 12-14) generally had larger brain mask volumes and greater inter-subject variability. These failure rates are comparable to those observed in studies of human scans.

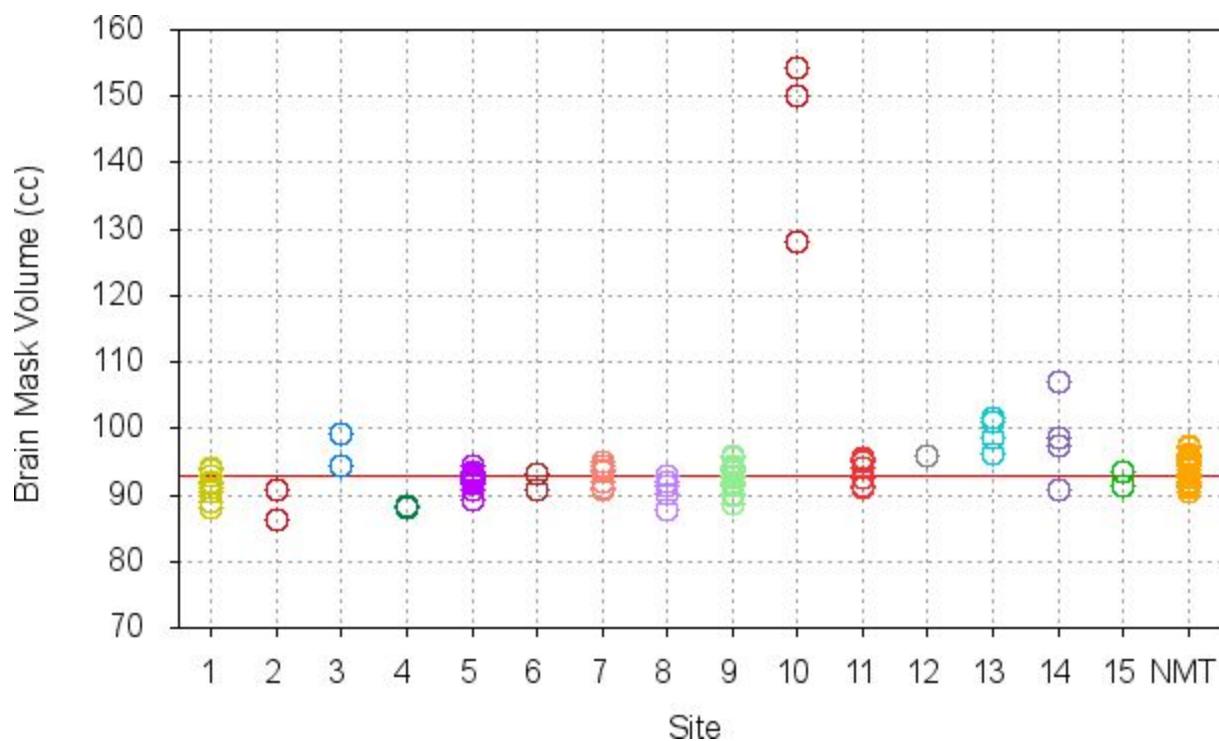


Figure 9: Brain mask volume in NMT standardized space for N=95 scans from the INDI PRIME-DE consortium and N=31 scans from the NMT cohort. The brain mask volume for the NMT volume is 92.8cc (red line). The small standard deviation for the volume of the brain mask in standardized space attests to the robustness of the brain extraction tool and the reliability of the volumetric registration to the NMT. The few obvious outliers show the processing failures.

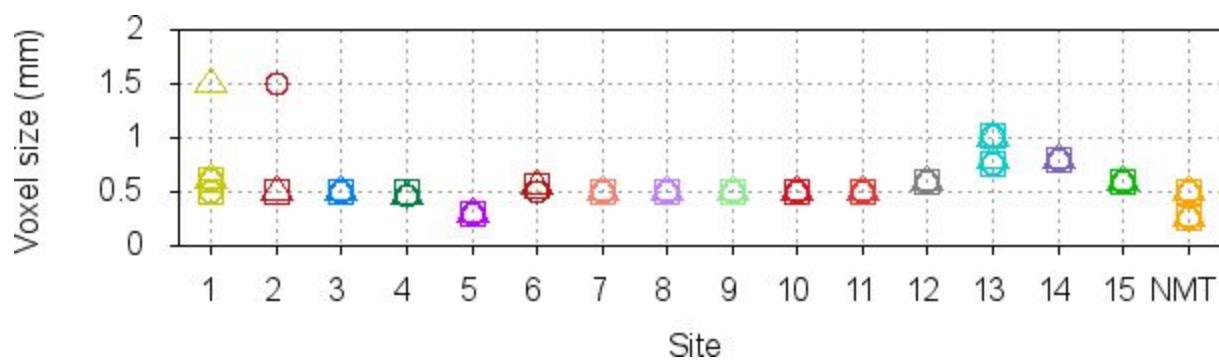


Figure 10: Acquisition voxel size for the N=95 scans from the INDI PRIME-DE consortium and N=31 scans from the NMT cohort. The three symbols (square, circle, triangle) refer to the voxel size in each of the 3 planes. By comparison to Figure 9, scans with a voxel size above 0.5mm tend to show higher deviation and variability from the volume of the brain mask of the NMT.

3.2 Cortical surface generation

Cortical surfaces were obtained for 29 of the 31 subjects in the NMT cohort (same two exclusions as for the creation of the average surface templates). The scans were processed at 0.25mm voxel size in stereotaxic space and the surfaces were produced at 40,962 vertices per hemisphere. A value of 75 was used as the spline parametric distance for N3. The *tlaplace* metric was used for the calculation of cortical thickness, blurred isotropically at a fwhm of 8mm. The disk usage per processed subject is about 325Mb (for short data type scans) to 550Mb (for float data type scans). (The T1w image input data type was preserved during processing.) Processing time ranged from 15-18 hours per subject. (For pre-processing only without surface extraction, the processing time is about 1.5 hours per subject.)

Figure 11 shows an overview of the CIVET stages and results for 4 NMT subjects. The results display the same level of accuracy for all subjects. Surface contours superimposed on the volume (top right panel for each subject) generally showed faithful correspondence to the gray and white borders. For some NMT subjects, the gray matter surface did not fully extend to the CSF on the orbital frontal cortex (OFC) surface, especially around the gyrus rectus (not shown), leading to artificially smaller OFC thickness in these cases. These errors are driven by segmentation errors due to scanning artifacts present for some of the NMT subjects and may not be an issue for other datasets. In many cases, the WM surface along the insula was slightly internal to the WM/GM border, lying along the lateral surface of the putamen, especially dorsally where the claustrum is very thin (see Fig. 11). These shifts in the position of the WM surface are caused by a relaxation of the WM surface intensity gradient correction adjacent to the claustrum. Cortical thickness values for the insula are likely slightly inflated in these cases. The WM surface in this region was, however, smooth and did not exhibit the ‘claustrum invagination problem’ (Autio et al. 2020). Nor were there issues with segmentation in such problem areas as the ventral and anterior claustrum or the temporal stem.

Sulcal variability across the 4 subjects can be observed on both the WM and pial surfaces (Fig. 11, bottom row). Cortical thickness also exhibited some variability across subjects. The 3 subjects shown in Koo et al. (2012) show similar sulcal and thickness variability. In general, thickness increased from posterior to anterior.

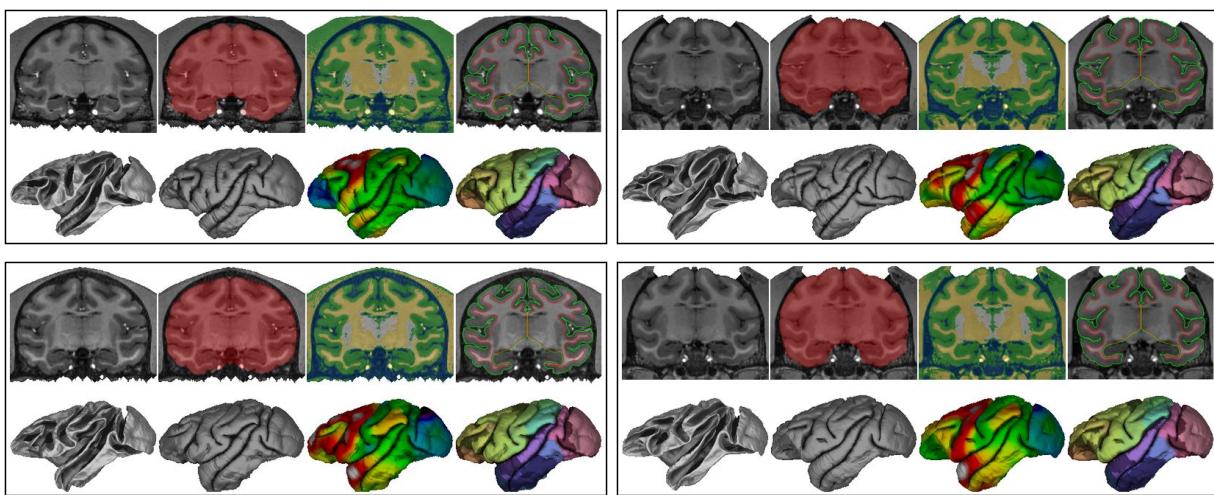


Figure 11: CIVET-macaque results for 4 NMT subjects. Within each set of images for a single subject, the top row, from left to right: coronal slice of the T1w image, brain mask, tissue classification (CSF in blue, cortical GM in green, portions of non-cortical GM in white, WM and other non-cortical GM in yellow), and cortical surfaces overlaid on the T1w image (WM surface in red, pial surface in green) bottom row, from left to right: white and pial surfaces; cortical thickness (range 1.0mm light blue to 3.0mm red) and the Level-4 CHARM parcellation viewed on the pial surface (52 ROIs).

3.2.1 Cortical thickness

The average cortical thickness map across the 29 subjects and its coefficient of variation (COV=standard deviation/mean) are shown in Figure 12. The cortical thickness map patterns follow closely those obtained from human MRIs processed with CIVET, with cortical thickness increasing from posterior to anterior, as well as published macaque cortical thickness maps (Autio et al 2020; Calebrese et al. 2015; Donahue et al. 2018; Koo et al. 2012). The COV was low in most regions. Upon thorough examination of the cortical surfaces, the higher COV in some regions likely reflects a mixture of intersubject variability and limitations in the surface extraction process, namely the frontal pole (scanning artifacts due to nasal cavity), the temporal pole (presence of the middle cerebral artery), and the occipital lobe (lower gray-white contrast, owing to the high myelin content of this region [Donahue et al. 2018]).

3.2.2 Regional analysis

Regional cortical thicknesses based on the CHARM atlas are reported in Table 3. Thickness is reported for hierarchical atlas levels 1, 2, and 3 (with 4, 17, and 36 regions, respectively) in the left and right hemispheres. Regional thickness was highly symmetric between the hemispheres.

Mean cortical thickness, within the brain and averaged across the 29 subjects, was 1.99 ± 0.43 mm. The frontal lobe was thickest and most varied in its thickness across vertices (2.33 ± 0.43 mm), followed by the temporal (2.12 ± 0.41 mm), parietal (1.93 ± 0.31 mm), and occipital lobes (1.67 ± 0.25 mm). At level 2, particularly thick regions include the motor cortex (2.62 ± 0.36 mm), including primary motor cortex, premotor cortex, and the dorsal aspect of the cingulate gyrus, and the temporal pole (2.49 ± 0.49 mm). Particularly thin regions include the somatosensory cortex (1.89 ± 0.30 mm), the core and belt regions of the auditory cortex (1.79 ± 0.36 mm), and primary visual cortex (1.66 ± 0.22 mm), especially within the parieto-occipital sulcus and on the ventral surface in the inferior occipital sulcus. Figure 13 shows the histogram for the distribution of cortical thickness for several individuals and for the averaged cortical thickness across all 29 subjects (without blurring).

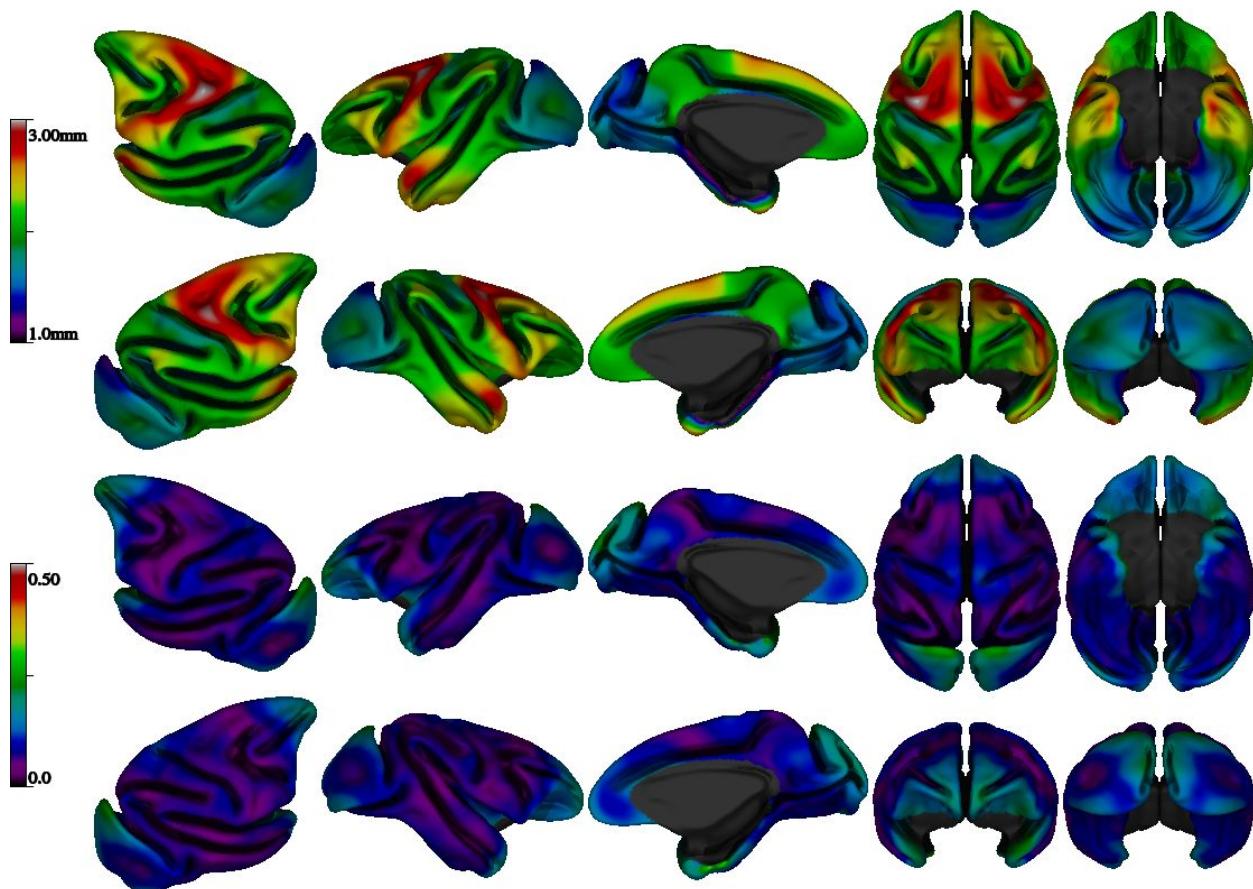


Figure 12: Mean cortical thickness (top) and coefficient of variation (COV) of cortical thickness (bottom) for $N=29$ subjects of the NMT dataset, using the *tlaplace* metric with *fwmm* of 8mm, viewed on the NMT average mid-cortical surface.

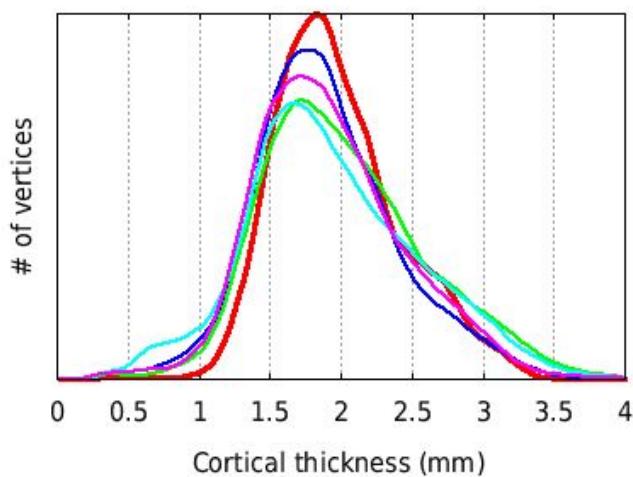


Figure 13: Histogram of mean cortical thickness for $N=29$ subjects of the NMT dataset (red thick line) and for 4 single NMT subjects (thin lines), using the *laplace* metric with no spatial smoothing ($fwhm = 0\text{mm}$). All 5 curves show the same total number of vertices (40,962 x 2 hemispheres) and area under the curve. The distribution for the mean curve is narrower (i.e., more central) than for most individual subjects. Consequently, the mean also has a higher peak than the other distributions.

CHARM parcellation			Cortical thickness (mm)	
Level	Label	Description	Left	Right
1	1	Frontal Lobe	2.33 ± 0.43	2.32 ± 0.43
2	2	anterior cingulate gyrus	1.99 ± 0.27	2.02 ± 0.30
3	3	<i>anterior cingulate cortex</i>	2.07 ± 0.29	$2.13. \pm 0.33$
3	11	<i>mid-cingulate cortex</i>	1.86 ± 0.14	1.85 ± 0.13
2	16	orbital frontal cortex	2.10 ± 0.33	2.10 ± 0.35
3	17	<i>medial orbital frontal cortex</i>	2.03 ± 0.27	2.07 ± 0.30
3	25	<i>lateral orbital frontal cortex</i>	2.10 ± 0.26	2.11 ± 0.27
3	37	<i>caudal orbital frontal cortex</i>	2.19 ± 0.51	2.13 ± 0.58
2	50	lateral prefrontal cortex	2.22 ± 0.36	2.20 ± 0.35
3	51	<i>periarcuate area 8A (Frontal Eye Fields)</i>	2.32 ± 0.19	2.31 ± 0.20
3	54	<i>dorsolateral prefrontal cortex</i>	2.27 ± 0.34	2.27 ± 0.33

3	64	<i>ventrolateral prefrontal cortex</i>	2.14 ± 0.38	2.10 ± 0.40
2	77	motor cortex	2.63 ± 0.36	2.61 ± 0.37
3	78	<i>lateral motor cortex</i>	2.66 ± 0.37	2.63 ± 0.37
3	87	<i>medial supplementary motor areas</i>	2.45 ± 0.22	2.40 ± 0.27
1	90	<u>Parietal Lobe</u>	1.94 ± 0.31	1.92 ± 0.31
2	91	somatosensory cortex	1.89 ± 0.30	1.88 ± 0.30
3	92	<i>primary somatosensory cortex</i>	1.87 ± 0.32	1.85 ± 0.32
3	95	<i>secondary somatosensory cortex</i>	1.98 ± 0.17	1.96 ± 0.18
2	96	superior parietal lobule	1.93 ± 0.23	1.90 ± 0.23
3	97	<i>visual areas V6 and V6A</i>	2.08 ± 0.17	2.08 ± 0.17
3	102	<i>area 5</i>	2.01 ± 0.20	1.99 ± 0.20
3	107	<i>ventromedial intraparietal sulcus</i>	1.76 ± 0.18	1.72 ± 0.16
2	112	inferior parietal lobule	2.01 ± 0.36	1.98 ± 0.36
3	113	<i>lateral intraparietal sulcus</i>	1.89 ± 0.19	1.85 ± 0.18
3	119	<i>medial superior temporal area</i>	1.72 ± 0.15	1.72 ± 0.13
3	120	<i>area 7 in inferior parietal lobule</i>	2.18 ± 0.40	2.15 ± 0.41
2	125	posterior medial cortex	1.97 ± 0.29	1.98 ± 0.29
3	126	<i>area 7 (PGm) on medial wall</i>	2.03 ± 0.20	2.05 ± 0.19
3	127	<i>posterior cingulate gyrus</i>	1.94 ± 0.33	1.95 ± 0.33
1	138	<u>Temporal Lobe</u>	2.13 ± 0.40	2.11 ± 0.42
2	139	medial temporal lobe	2.10 ± 0.66	1.99 ± 0.70
3	148	<i>parahippocampal cortex</i>	1.64 ± 0.25	1.52 ± 0.25
3	153	<i>rhinal cortex</i>	2.64 ± 0.57	2.56 ± 0.64
2/3	165	temporal pole	2.47 ± 0.51	2.51 ± 0.48
2	175	inferior temporal cortex	2.13 ± 0.23	2.13 ± 0.23
3	176	<i>area TEO</i>	2.09 ± 0.12	2.10 ± 0.11

3	177	area <i>TE</i>	2.25 ± 0.21	2.24 ± 0.22
3	188	<i>fundus of superior temporal sulcus</i>	1.92 ± 0.12	1.92 ± 0.12
2	193	superior temporal region	2.30 ± 0.33	2.28 ± 0.34
3	194	<i>rostral superior temporal region</i>	2.33 ± 0.36	2.32 ± 0.36
3	199	<i>caudal superior temporal gyrus</i>	2.23 ± 0.25	2.20 ± 0.26
2	204	core and belt areas of auditory cortex	1.80 ± 0.36	1.78 ± 0.36
3	205	<i>belt areas of auditory cortex</i>	1.78 ± 0.37	1.75 ± 0.34
3	218	<i>polar rostrotemporal cortex</i>	2.54 ± 0.45	2.64 ± 0.46
3	219	<i>core areas of auditory cortex</i>	1.74 ± 0.21	1.74 ± 0.23
2/3	224	floor of the lateral sulcus	2.04 ± 0.47	1.99 ± 0.53
1	231	<u>Occipital Lobe</u>	1.67 ± 0.25	1.66 ± 0.25
2/3	232	middle temporal area	1.72 ± 0.09	1.71 ± 0.08
2	233	extrastrate visual areas 2-4	1.67 ± 0.27	1.66 ± 0.27
3	234	<i>visual area 4</i>	1.91 ± 0.22	1.91 ± 0.21
3	237	<i>preoccipital visual areas 2-3</i>	1.61 ± 0.24	1.60 ± 0.25
2/3	246	primary visual cortex	1.67 ± 0.23	1.65 ± 0.21
		Total	1.99 ± 0.43	1.98 ± 0.44

Table 3: Regional mean cortical thickness of the NMT cohort (N=29) for the 3 broadest levels of the CHARM anatomical atlas, sorted by ordinal label number. Region names and atlas number labels are listed in different fonts and with different degrees of indentation depending on the CHARM level(s) in which they appear. Levels 1, 2, and 3, which contain 4, 17, and 36 regions, respectively, are included. For each subject, the mean and standard deviation of the thickness (tlaplace metric with fwhm of 0mm) was computed across the vertices within each cortical region and hemisphere. Each region's mean and standard deviation was then averaged across subjects. The overall standard deviation is higher than the standard deviation of individual specialized regions, indicating higher inter-regional variability than intra-regional variability. Thinner areas include somatosensory, auditory and visual cortices, while temporal pole and motor cortex were the thickest areas. The thickness of each region was nearly identical between the left and right hemispheres. The smaller cortical thickness values in the parahippocampal

cortex (region 148) reflect its proximity to the hippocampus, which is not modelled. This region should be excluded in analyses.

3.2.3 Surface area

The total area of the white matter and pial surfaces is displayed in Figure 14 for the NMT subjects (N=29). This total area includes the mid-plane wall cutting through the corpus callosum and the brainstem. Despite the inter-subject variability, areas of the left and right hemispheres are in close agreement within each subject for both cortical surfaces. The mean area per hemisphere was $10,098 \text{ mm}^2$ for the white matter surface and $13,441 \text{ mm}^2$ for the pial surface. These values are larger than those reported for the macaque (Autio et al 2020; Donahue et al. 2018; Koo et al. 2012), even when accounting for the mid-plane wall region. The larger values obtained here could be an indication of deeper sulcal penetration of the pial surface and increased morphological details captured by the extracted surfaces.

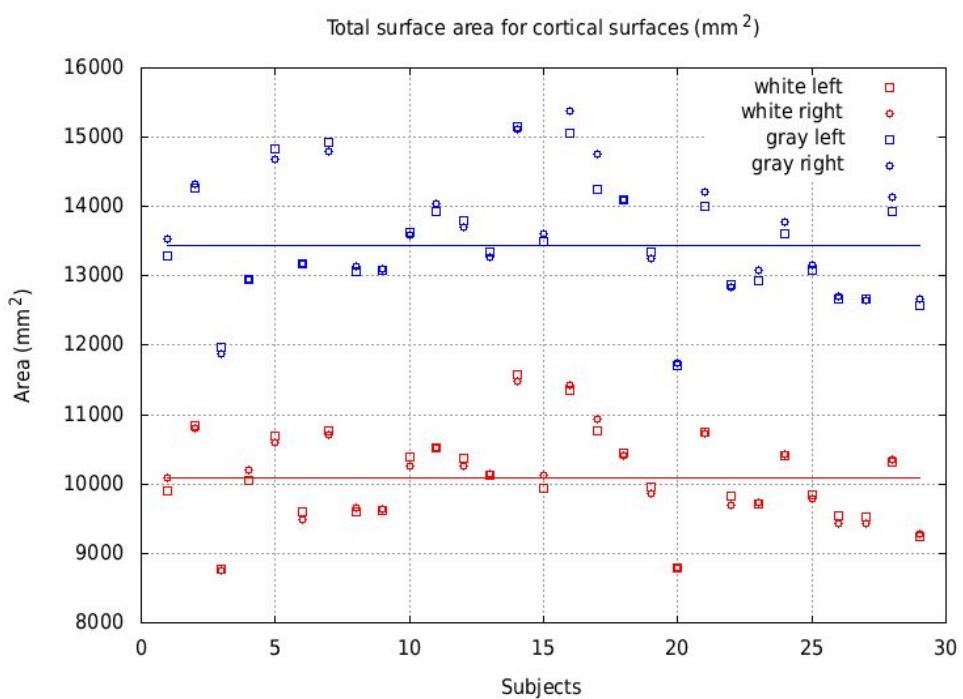


Figure 14: Total area for the white matter (red) and pial (blue) surfaces for the NMT cohort (N=29), for left and right hemispheres. The area per hemisphere, averaged across subjects and hemispheres, is $10,098 \text{ mm}^2$ for the white matter surface (red line) and $13,441 \text{ mm}^2$ for the pial surface (blue line). Left and right surface areas agreed for all subjects, independently of inter-subject variability.

4. Discussion

We have introduced CIVET-macaque, a pipeline for automated reconstruction of cortical surfaces of macaque brains. The pipeline requires only a T1-weighted MRI image to generate high resolution cortical surfaces, capable of intersubject surface-based coregistration with per-vertex correspondence between gray and white matter surfaces. Preprocessing and surface reconstruction were customized for macaques to address species related issues such as variations in head size and brain morphology.

The pipeline has been developed primarily on 29 scans out of the 31 subjects in the NMT cohort. Further tests were conducted on a larger heterogeneous collection of 95 scans from the INDI PRIME-DE consortium. The success rate depended, inevitably, on the quality of the imaging data. Input voxel resolution, field inhomogeneities, the presence of visible blood vessels, and scanning artifacts can all potentially cause various defects in the cortical surfaces. Supplementary Figure 1 shows the typical variability amongst scans and some common artifacts and their repercussion on the outcome of the pipeline. Some scans have a desirable high GM-WM contrast, but suffer from scanning artifacts anterior to the frontal and temporal poles. These artifacts interfered with surface placement, in particular for the NMT cohort, causing the COV for cortical thickness in these regions to peak. Some other scans were mostly free of scanning artifacts, but had low GM-WM contrast, which proved problematic during the extraction of the white matter surface (intensity gradient correction). Still other scans were of acceptable quality and produced expected results. It is hoped that scanning protocols better suited to corticometry analyses will be employed in the future for additional structural data acquisition in monkeys.

Overall, the pre-processing phases of CIVET-macaque proved to be highly reliable and robust with regard to registration and brain-masking, independently of the quality of the scans. Cortical surface extraction was also successful although more challenging, being very sensitive to the quality of the data and, in particular, to the presence of high intensities representing blood vessels and other acquisition artifacts. The availability of a T2w scan could be used, in the future, to improve the detection of such blood vessels. Finally, sufficient scan resolution (0.5mm voxel size or smaller) is necessary to delineate the fine structures of the macaque brain

morphology, unlike the human brain for which a scan resolution at 1mm voxel size has been employed for many years.

At present, to circumvent possible shortcomings in the white matter surface extraction, a brain mask with special annotations can be supplied with the input T1w scan, in native space. This initial custom brain mask is obtained from a first run of the pipeline and annotated in problematic areas where errors occurred during the white matter surface extraction. Label=1 represents the underlying brain mask. Label=2 represents offending misclassified voxels that should not be white matter (e.g. blood vessels, meninges/dura, gray matter with high signal intensities). Label=3 represents missing white matter voxels (e.g. reconstructed thin parahippocampal gyrus or incompletely filled ventricles for the purpose of white matter surface extraction). Basically, the annotated mask allows construction of a valid white matter mask for the white matter surface extraction by retaining classified WM within the cerebrum (label ≥ 1), while removing voxels incorrectly labeled as WM (label = 2), and adding missing WM voxels to preserve the spherical topology of the cortical mantle (label = 3). Hollow ventricles can also be closed completely (label = 3) to avoid leakage of the white matter surface into the ventricles. CIVET is restarted with the new annotated custom mask and new surfaces are obtained.

All 31 NMT T1w scans revealed high-intensity voxels representing blood vessels or other scanning artifacts causing various white matter surface extraction errors of different degrees. An annotated mask, in native space, was thus needed for each of the 29 NMT scans included in this study. The task was fairly quick and amenable for this small dataset. After the masks were annotated, the pipeline was restarted in a fully automated way. The ensuing white matter surfaces were of high quality, as specifically needed for the generation of the average white matter and mid-cortical surfaces with minimal variability due to algorithmic noise.

The generated maps of cortical thickness for the NMT cohort (Figure 12 and Table 3) revealed expected morphological trends. These include thinner somatosensory, visual and auditory areas, thicker temporal pole and motor cortex, with the remaining homotypic isocortex exhibiting thicknesses in between these extremes. This pattern is consistent with existing studies of macaque MRI morphology (Wagstyl et al. 2015) and similar to patterns seen in human histological studies of cortical thickness in the human BigBrain (Amunts et al. 2013; Wagstyl et

al. 2020). CIVET-macaque would be an essential step towards creating an equivalent 3D histological atlas of the macaque brain.

The thickness map in Fig. 12 is qualitatively similar to that of published maps of individual subjects (see Koo et al. 2012; Oguz et al. 2015; Fig. 4 of Autio et al. 2020) and group averages (see Fig. 3 of Koo et al. 2012, n=8; Fig. 4 of Autio et al. 2020, n=12; Fig. S2 of Seidlitz et al. 2017, n=31), in particular that in Fig. S4 of Donahue et al. (2018, n=19) and Fig. 6 of Calebrese et al. (2015, n=10). The mean cortical thickness ($t_{laplace}$, $fw\text{hm}=8\text{mm}$) reported here is $1.99 \pm 0.17\text{ mm}$. The mean thickness of the NMT cohort agrees with that reported by Donahue et al. (2018, $2.0 \pm 0.1\text{ mm}$, n=19) and Autio et al (2020, $2.1 \pm 0.1\text{ mm}$, n=4), but is smaller than the value reported by Koo et al. (2012, $2.37 \pm 0.19\text{ mm}$, n=18). The pipeline employed by Koo et al. is a variant from an older version of CIVET. Their surfaces included on the order of 5000 vertices and the t_{link} metric was used for cortical thickness, which leads to an overestimation due to mesh distortion during the expansion of the pial surface. Note that the standard deviation, here, is the vertex-wise standard deviation across the n subjects, relative to the local vertex-wise mean of cortical thickness for the n subjects, averaged across the number of vertices on the surface (by hemisphere). It reflects how much cortical thickness varies locally - at the vertex level - across the subjects, relative to the local mean.

The variance in cortical thickness across the NMT cohort was low in most places. Greater COV (see Fig. 12) in the frontal pole, superior part of the occipital lobe and temporal pole appears in part to be driven by poor scan quality and algorithmic shortcomings. The frontal pole suffers from scanning artifacts due to the nasal cavity. In those subjects with the thinnest occipital lobe, the gray-white contrast was poor in that region, causing the true gray-white intensity transition to be missed and the white matter surface to be displaced towards the pial boundary. Finally, blood vessels (namely, the middle cerebral artery) in the vicinity of the temporal pole appear to be the source of the defects in both the white and gray matter surfaces in that region. Overall, the regions of high COV in this cohort are conjectured to be driven by processing errors and to be a weak indicator of true biological variance. Caution should be used when reporting results of biological relevance in these regions.

5. Conclusions

CIVET-macaque is the first true fully automated pipeline for processing macaque brains for corticometric studies. The pipeline, based on the imaging minc data format, has been developed in PERL and C/C++ for Linux systems. The open-source project (https://github.com/aces/CIVET_Full_Project) is also found on the PRIME-Resource Exchange portal (<https://prime-re.github.io/>). Outputs are produced in minc (HDF5-based) and NIfTI/GIfTI formats.

This new software enables the field of NHP imaging to further explore questions of structure and function across large populations and datasets in a new way. It is anticipated that the open usage of CIVET-macaque will promote collaborative efforts in data collection, sharing, and automated analyses from which the non-human primate brain-image field will further advance (Autio et al. 2020). For example, corticometry would provide a natural and complementary insight on longitudinal developmental studies in the macaque monkey (Scott et al. 2016; Wang et al. 2018). Finally, it is our intent to share complete CIVET outputs for the INDI PRIME-DE on the PRIME-RE portal (<https://prime-re.github.io/>) (Messinger et al., *this issue*).

CIVET has the potential to be expanded to other non-human primate species, with possible extensions to the marmoset and the chimpanzee, using repositories of previously collected and, if possible, publically available anatomical scans. The ability to process marmoset, macaque, chimpanzee, and human data with CIVET will allow for comparative studies of surface-based morphometry.

Data accessibility and availability

Scans used for the development of CIVET-macaque are already publicly accessible. These include the D99 post-mortem brain and atlas (<https://afni.nimh.nih.gov/Macaque>), the NMT volume (version 1.2; <https://afni.nimh.nih.gov/NMT>), and single subject scans from the INDI PRIME-DE repository (http://fcon_1000.projects.nitrc.org/indi/indiPRIME.html/). CIVET-macaque will be versioned and added to the global open-source CIVET project (https://github.com/aces/CIVET_Full_Project). Binaries for common Linux platforms will be provided for convenience. CIVET-macaque, the NMT, the CHARM (Jung et al., *this issue*), and updates to these resources can be centrally accessed from the PRIME-Resource Exchange portal (<https://prime-re.github.io/>) (Messinger et al., *this issue*).

CRediT authorship contribution statement

Claude Lepage: Conceptualization, software, data processing, formal analysis, writing. **Konrad Wagstyl:** Conceptualization, writing. **Benjamin Jung:** Resources, writing, software. **Jakob Seidlitz:** Conceptualization, resources, writing. **Caleb Sponheim:** Conceptualization, resources, writing. **Leslie Ungerleider:** Supervision, funding acquisition. **Xindi Wang:** Software and data curating, testing. **Alan C. Evans:** Supervision, funding acquisition. **Adam Messinger:** Conceptualization, resources, investigation, writing, validation, visualization, project administration, supervision.

Declaration of competing interest

The authors report no competing interest.

Acknowledgments

This research was supported (in part) by the Intramural Research Program of the NIMH and includes the relevant Annual Report number in the following format (ZICMH002899). Parts of this work have received support from Healthy Brains for Healthy Lives. KW is supported by the Wellcome Trust (215901/Z/19/Z). The authors report no conflicts of interest.

References

Amunts, Katrin, Claude Lepage, Louis Borgeat, Hartmut Mohlberg, Timo Dickscheid, Marc-Étienne Rousseau, Sebastian Bludau, et al. 2013. “BigBrain: An Ultrahigh-Resolution 3D Human Brain Model.” *Science* 340 (6139): 1472–75.

Autio, Joonas A., Matthew F. Glasser, Takayuki Ose, Chad J. Donahue, Matteo Bastiani, Masahiro Ohno, Yoshihiko Kawabata, et al. 2020. “Towards HCP-Style Macaque Connectomes: 24-Channel 3T Multi-Array Coil, MRI Sequences and Preprocessing.” *NeuroImage*, April, 116800.

Avants, Brian B., Nicholas J. Tustison, Gang Song, Philip A. Cook, Arno Klein, and James C. Gee. 2011. “A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration.” *NeuroImage* 54 (3): 2033–44.

Avants, Brian B., Paul Yushkevich, John Pluta, David Minkoff, Marc Korczykowski, John Detre, and James C. Gee. 2010. “The Optimal Template Effect in Hippocampus Studies of Diseased Populations.” *NeuroImage* 49 (3): 2457–66.

Bakken, Trygve E., Jeremy A. Miller, Song-Lin Ding, Susan M. Sunkin, Kimberly A. Smith, Lydia Ng, Aaron Szafer, et al. 2016. “A Comprehensive Transcriptional Map of Primate Brain Development.” *Nature* 535 (7612): 367–75.

Bonaiuto, James J., Holly E. Rossiter, Sofie S. Meyer, Natalie Adams, Simon Little, Martina F. Callaghan, Fred Dick, Sven Bestmann, and Gareth R. Barnes. 2018. “Non-Invasive Laminar Inference with MEG: Comparison of Methods and Source Inversion Algorithms.”

NeuroImage 167 (February): 372–83.

Boucher, Maxime, Sue Whitesides, and Alan Evans. 2009. “Depth Potential Function for Folding Pattern Representation, Registration and Analysis.” *Medical Image Analysis* 13 (2): 203–14.

Calabrese, Evan, Alexandra Badea, Christopher L. Coe, Gabriele R. Lubach, Yundi Shi, Martin A. Styner, and G. Allan Johnson. 2015. “A Diffusion Tensor MRI Atlas of the Postmortem Rhesus Macaque Brain.” *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2015.05.072>.

Collins, D. L., P. Neelin, T. M. Peters, and A. C. Evans. 1994. “Automatic 3D Intersubject Registration of MR Volumetric Data in Standardized Talairach Space.” *Journal of Computer Assisted Tomography* 18 (2): 192–205.

Deichmann, R., C. Schwarzbauer, and R. Turner. 2004. “Optimisation of the 3D MDEFT Sequence for Anatomical Brain Imaging: Technical Implications at 1.5 and 3 T.” *NeuroImage* 21 (2): 757–67.

Donahue, Chad J., Matthew F. Glasser, Todd M. Preuss, James K. Rilling, and David C. Van Essen. 2018. “Quantitative Assessment of Prefrontal Cortex in Humans Relative to Nonhuman Primates.” *Proceedings of the National Academy of Sciences of the United States of America* 115 (22): E5183–92.

Fischl, Bruce. 2012. “FreeSurfer.” *NeuroImage* 62 (2): 774–81.

Fischl, B., M. I. Sereno, R. B. Tootell, and A. M. Dale. 1999. “High-Resolution Intersubject Averaging and a Coordinate System for the Cortical Surface.” *Human Brain Mapping* 8 (4): 272–84.

Glasser, Matthew F., Timothy S. Coalson, Emma C. Robinson, Carl D. Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, et al. 2016. “A Multi-Modal Parcellation of Human Cerebral Cortex.” *Nature* 536 (7615): 171–78.

Gorgolewski, Krzysztof J., Gael Varoquaux, Gabriel Rivera, Yannick Schwarz, Satrajit S. Ghosh, Camille Maumet, Vanessa V. Sochat, et al. 2015. “NeuroVault.org: A Web-Based Repository for Collecting and Sharing Unthresholded Statistical Maps of the Human Brain.” *Frontiers in Neuroinformatics* 9 (April): 8.

Greve, Douglas N., Claus Svarer, Patrick M. Fisher, Ling Feng, Adam E. Hansen, William Baare, Bruce Rosen, Bruce Fischl, and Gitte M. Knudsen. 2014. “Cortical Surface-Based Analysis Reduces Bias and Variance in Kinetic Modeling of Brain PET Data.” *NeuroImage* 92 (May): 225–36.

Hawrylycz, Michael J., Ed S. Lein, Angela L. Guillozet-Bongaarts, Elaine H. Shen, Lydia Ng, Jeremy A. Miller, Louie N. van de Lagemaat, et al. 2012. “An Anatomically Comprehensive Atlas of the Adult Human Brain Transcriptome.” *Nature* 489 (7416): 391–99.

Horsley, Victor, and R. H. Clarke. 1908. “THE STRUCTURE AND FUNCTIONS OF THE CEREBELLUM EXAMINED BY A NEW METHOD.” *Brain*. <https://doi.org/10.1093/brain/31.1.45>.

Huth, Alexander G., Wendy A. de Heer, Thomas L. Griffiths, Frédéric E. Theunissen, and Jack L. Gallant. 2016. “Natural Speech Reveals the Semantic Maps That Tile Human Cerebral Cortex.” *Nature* 532 (7600): 453–58.

Jung, Benjamin, Paul A. Taylor, Jakob Seidlitz, Caleb Sponheim, Daniel Glen, and Adam Messinger. “A comprehensive macaque fMRI pipeline and hierarchical atlas.” *NeuroImage*. *this issue*.

Kim, June Sic, Vivek Singh, Jun Ki Lee, Jason Lerch, Yasser Ad-Dab’bagh, David MacDonald, Jong Min Lee, Sun I. Kim, and Alan C. Evans. 2005. “Automated 3-D Extraction and Evaluation of the Inner and Outer Cortical Surfaces Using a Laplacian Map and Partial Volume Effect Classification.” *NeuroImage* 27 (1): 210–21.

Koo, Bang-Bon, Steven P. Schettler, Donna E. Murray, Jong-Min Lee, Ronald J. Killiany, Douglas L. Rosene, Dae-Shik Kim, and Itamar Ronen. 2012. "Age-Related Effects on Cortical Thickness Patterns of the Rhesus Monkey Brain." *Neurobiology of Aging* 33 (1): 200.e23–31.

Lee, Jing-Huei, Michael Garwood, Ravi Menon, Gregor Adriany, Peter Andersen, Charles L. Truwit, and Kämil Uğurbil. 1995. "High Contrast and Fast Three-Dimensional Magnetic Resonance Imaging at High Fields." *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 34 (3): 308–12.

Lepage, Claude, Lindsay Lewis, Seun Jeun, Patrick Bermudez, Najmeh Khalili-Mahani, Mona Omidyegaheh, Alex Zijdenbos, Robert D. Vincent, Reza Adalat, and Alan C. Evans. 2017. "Human MR Evaluation of Cortical Thickness Using CIVET v2.1." *Organization for Human Brain Mapping*. <https://ww5.aievolution.com/hbm1701/index.cfm?do=doc.viewDocument&documentID=1915>.

Lerch, Jason P., and Alan C. Evans. 2005. "Cortical Thickness Analysis Examined through Power Analysis and a Population Simulation." *NeuroImage* 24 (1): 163–73.

Lerch, Jason P., Jens C. Pruessner, Alex Zijdenbos, Harald Hampel, Stefan J. Teipel, and Alan C. Evans. 2005. "Focal Decline of Cortical Thickness in Alzheimer's Disease Identified by Computational Neuroanatomy." *Cerebral Cortex* 15 (7): 995–1001.

Love, Scott A., Damien Marie, Muriel Roth, Romain Lacoste, Bruno Nazarian, Alice Bertello, Olivier Coulon, Jean-Luc Anton, and Adrien Meguerditchian. 2016. "The Average Baboon Brain: MRI Templates and Tissue Probability Maps from 89 Individuals." *NeuroImage* 132 (May): 526–33.

Lyttelton, Oliver, Maxime Boucher, Steven Robbins, and Alan Evans. 2007. "An Unbiased Iterative Group Registration Template for Cortical Surface Analysis." *NeuroImage* 34 (4): 1535–44.

Messinger, Adam, Ting Xu, Xindi Wang, Régis Trapeau, Roberto Toro, Paul Taylor, Caleb Sponheim, et al. "An open resource for non-human primate neuroimaging data analysis." *NeuroImage*. this issue.

Milham, Michael P., Lei Ai, Bonhwang Koo, Ting Xu, Céline Amiez, Fabien Balezeau, Mark G. Baxter, et al. 2018. "An Open Resource for Non-Human Primate Imaging." *Neuron* 100 (1): 61–74.e2.

Miranda-Dominguez, Oscar, Brian D. Mills, David Grayson, Andrew Woodall, Kathleen A. Grant, Christopher D. Kroenke, and Damien A. Fair. 2014. "Bridging the Gap between the Human and Macaque Connectome: A Quantitative Comparison of Global Interspecies Structure-Function Relationships and Network Topology." *Journal of Neuroscience* 34 (16): 5552–63.

Oguz, Ipek, Martin Styner, Mar Sanchez, Yundi Shi, and Milan Sonka. 2015. "LOGISMOS-B for Primates: Primate Cortical Surface Reconstruction and Thickness Measurement." *Proceedings of SPIE The International Society for Optical Engineering* 9413. <https://doi.org/10.1117/12.2082327>.

Reveley, Colin, Audrunas Gruslys, Frank Q. Ye, Daniel Glen, Jason Samaha, Brian E Russ, Ziad Saad, Anil K Seth, David A. Leopold, and Kadharbatcha S. Saleem. 2017. "Three-Dimensional Digital Template Atlas of the Macaque Brain." *Cerebral Cortex* 27 (9): 4463–77.

Robbins, S. 2003. "Anatomical Standardization of the Human Brain in Euclidean 3-Space and

on the Cortical 2-Manifold [PhD Thesis]." *School of Computer Science, McGill University, Montreal, Quebec, Canada.*

Robinson, Emma C., Saad Jbabdi, Matthew F. Glasser, Jesper Andersson, Gregory C. Burgess, Michael P. Harms, Stephen M. Smith, David C. Van Essen, and Mark Jenkinson. 2014. "MSM: A New Flexible Framework for Multimodal Surface Matching." *NeuroImage* 100 (October): 414–26.

Saleem, Kadharbatcha S., and Nikos K. Logothetis. 2012. *A Combined MRI and Histology Atlas of the Rhesus Monkey Brain in Stereotaxic Coordinates*. Academic Press.

Scott, Julia A., David Grayson, Evan Fletcher, Aaron Lee, Melissa D. Bauman, Cynthia Mills Schumann, Michael H. Buonocore, and David G. Amaral. 2016. "Longitudinal Analysis of the Developing Rhesus Monkey Brain Using Magnetic Resonance Imaging: Birth to Adulthood." *Brain Structure and Function*. <https://doi.org/10.1007/s00429-015-1076-x>.

Seidlitz, Jakob, Caleb Sponheim, Daniel Glen, Frank Q. Ye, Kadharbatcha S. Saleem, David A. Leopold, Leslie Ungerleider, and Adam Messinger. 2018. "A Population MRI Brain Template and Analysis Tools for the Macaque." *NeuroImage* 170 (April): 121–31.

Sherif, Tarek, Pierre Rioux, Marc-Etienne Rousseau, Nicolas Kassis, Natacha Beck, Reza Adalat, Samir Das, Tristan Glatard, Alan C. Evans. 2014. "CBRAIN: A web-based, distributed computing platform for collaborative neuroimaging research." *Front Neurosci* 8(54):1-13.

Sled, J. G., A. P. Zijdenbos, and A. C. Evans. 1998. "A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data." *IEEE Transactions on Medical Imaging* 17 (1): 87–97.

Smith, Stephen M. 2002. "Fast Robust Automated Brain Extraction." *Human Brain Mapping* 17 (3): 143–55.

The HDF Group. 2016. "Hierarchical Data Format, Version 5, 1997--2016." URL: [Http://www.hdfgroup.org/HDF5](http://www.hdfgroup.org/HDF5).

Tohka, Jussi, Alex Zijdenbos, and Alan Evans. 2004. "Fast and Robust Parameter Estimation for Statistical Partial Volume Models in Brain MRI." *NeuroImage* 23 (1): 84–97.

Tustison, N.J., P.A. Cook, A. Klein, G. Song, S.R. Das, J.T. Duda, B.M. Kandel, N. van Strien, J.R. Stone, J.C. Gee, and B.B. Avants. 2014. "Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements." *NeuroImage* 99:166–179.

Vincent, Robert D., Peter Neelin, Najmeh Khalili-Mahani, Andrew L. Janke, Vladimir S. Fonov, Steven M. Robbins, Leila Baghdadi, et al. 2016. "MINC 2.0: A Flexible Format for Multi-Modal Images." *Frontiers in Neuroinformatics* 10 (August): 35.

Wagstyl, Konrad, Stéphanie Larocque, Guillem Cucurull, Claude Lepage, Joseph Paul Cohen, Sebastian Bludau, Nicola Palomero-Gallagher, et al. 2020. "BigBrain 3D Atlas of Cortical Layers: Cortical and Laminar Thickness Gradients Diverge in Sensory and Motor Cortices." *PLoS Biology* 18 (4): e3000678.

Wagstyl, Konrad, Lisa Ronan, Ian M. Goodyer, and Paul C. Fletcher. 2015. "Cortical Thickness Gradients in Structural Hierarchies." *NeuroImage* 111 (C): 241–50.

Wang, F., C. Lian, J. Xia, Z. Wu, D. Duan, L. Wang, D. Shen, and G. Li. 2018. "Construction of Spatiotemporal Infant Cortical Surface Atlas of Rhesus Macaque." In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 704–7.

Weiskopf, Nikolaus, John Suckling, Guy Williams, Marta M. Correia, Becky Inkster, Roger Tait, Cinly Ooi, Edward T. Bullmore, and Antoine Lutti. 2013. "Quantitative Multi-Parameter Mapping of R1, PD(*), MT, and R2(*) at 3T: A Multi-Center Validation." *Frontiers in Neuroscience* 7 (June): 95.

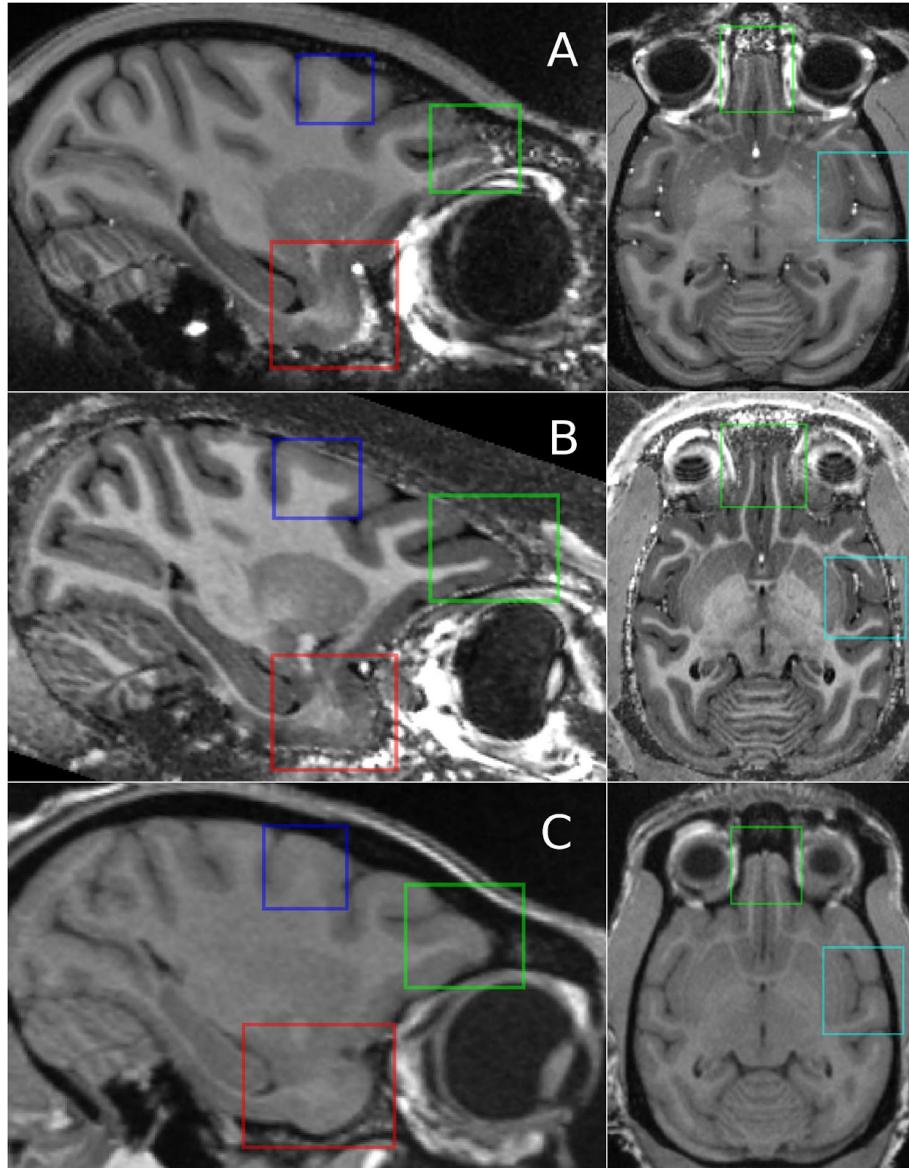
Xu, Ting, Karl-Heinz Nenning, Ernst Schwartz, Seok-Jun Hong, Joshua T. Vogelstein, Damien

A. Fair, Charles E. Schroeder, et al. 2019. "Cross-Species Functional Alignment Reveals Evolutionary Hierarchy Within the Connectome." *bioRxiv*. <https://doi.org/10.1101/692616>.

Yeo, B. T. Thomas, Fenna M. Krienen, Jorge Sepulcre, Mert R. Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L. Roffman, et al. 2011. "The Organization of the Human Cerebral Cortex Estimated by Intrinsic Functional Connectivity." *Journal of Neurophysiology* 106 (3): 1125–65.

Young, Jeffrey T., Yundi Shi, Marc Niethammer, Michael Grauer, Christopher L. Coe, Gabriele R. Lubach, Bradley Davis, et al. 2017. "The UNC-Wisconsin Rhesus Macaque Neurodevelopment Database: A Structural MRI and DTI Database of Early Postnatal Development." *Frontiers in Neuroscience* 11 (February): 29.

Zijdenbos, Alex, Reza Forghani, and Alan Evans. 1998. "Automatic Quantification of MS Lesions in 3D MRI Brain Data Sets: Validation of INSECT." In *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, edited by William M. Wells, Alan Colchester, and Scott Delp, 1496:439–48. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg.



Supplementary Figure 1: Scan variability and problem areas for 3 subjects from different sites (A: NMT subject, NIMH; B: PRIME-DE subject, site 10; C: PRIME-DE subject, site 11). GM-WM contrast (blue boxes) was clear for subjects A and B but very weak for subject C. Conversely, hyperintensities around the temporal pole (red boxes) and artifacts anterior to the frontal pole (green boxes), which likely arise due to proximity to the sinuses, were present for subjects A and B but not subject C. These hyperintensities diffused into the cortex causing local errors in the placement of the white matter surface. The cyan boxes reveal the presence of hyperintense voxels for blood vessels in subjects A and B (errors in white matter surface extraction), the weak GM-WM delineation of the claustrum in all subjects (misplaced white matter surface), and the presence of high intensity voxels for the bone marrow within the skull in subject B (leakage of brain mask into scalp).