

# The impact of the Siemens Tim Trio to Prisma upgrade and the addition of volumetric navigators on cortical thickness, structure volume, and <sup>1</sup>H-MRS indices: An MRI reliability study with implications for longitudinal study designs

Eric Plitman<sup>a,b,\*</sup>, Aurélie Bussy<sup>a,c</sup>, Vanessa Valiquette<sup>a</sup>, Alyssa Salaciak<sup>a</sup>, Raihaan Patel<sup>a,d</sup>, Lani Cupo<sup>a,c</sup>, Marie-Lise Béland<sup>a</sup>, Stephanie Tullo<sup>a,c</sup>, Christine Lucas Tardif<sup>d,e,f</sup>, M. Natasha Rajah<sup>b,g,h</sup>, Jamie Near<sup>b,d,h</sup>, Gabriel A. Devenyi<sup>a,b,h</sup>, M. Mallar Chakravarty<sup>a,b,d,h,\*</sup>

<sup>a</sup> Computational Brain Anatomy Laboratory, Cerebral Imaging Centre, Douglas Mental Health University Institute, Montreal, Quebec, Canada

<sup>b</sup> Department of Psychiatry, McGill University, Montreal, Quebec, Canada

<sup>c</sup> Integrated Program in Neuroscience, McGill University, Montreal, Canada

<sup>d</sup> Department of Biomedical Engineering, McGill University, Montreal, Quebec, Canada

<sup>e</sup> McConnell Brain Imaging Centre, Montreal Neurological Institute, Quebec, Canada

<sup>f</sup> Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada

<sup>g</sup> Department of Psychology, McGill University, Montreal, Quebec, Canada

<sup>h</sup> Cerebral Imaging Centre, Douglas Mental Health University Institute, Montreal, Quebec, Canada

## ARTICLE INFO

### Keywords:

Siemens

Prisma

Volumetric Navigators

Cortical thickness

Volume

Proton magnetic resonance spectroscopy

## ABSTRACT

Many magnetic resonance imaging (MRI) measures are being studied longitudinally to explore topics such as biomarker detection and clinical staging. A pertinent concern to longitudinal work is MRI scanner upgrades. When upgrades occur during the course of a longitudinal MRI neuroimaging investigation, there may be an impact on the compatibility of pre- and post-upgrade measures. Similarly, subject motion is another issue that may be detrimental to MRI work and embedding volumetric navigators (vNavs) within acquisition sequences has emerged as a technique that allows for prospective motion correction. Our research group recently underwent an upgrade from a Siemens MAGNETOM 3T Tim Trio system to a Siemens MAGNETOM 3T Prisma Fit system. The goals of the current work were to: 1) investigate the impact of this upgrade on commonly used structural imaging measures and proton magnetic resonance spectroscopy indices (“Prisma Upgrade protocol”) and 2) examine structural imaging measures in a sequence with vNavs alongside a standard acquisition sequence (“vNav protocol”). While high reliability was observed for most of the investigated MRI outputs, suboptimal reliability was observed for certain indices. Across the scanner upgrade, increases in frontal, temporal, and cingulate cortical thickness (CT) and thalamus volume, along with decreases in parietal CT and amygdala, globus pallidus, hippocampus, and striatum volumes, were observed. No significant impact of the upgrade was found in <sup>1</sup>H-MRS analyses. Further, CT estimates were found to be larger in MPRAGE acquisitions compared to vNav-MPRAGE acquisitions mainly within temporal areas, while the opposite was found mostly in parietal brain regions. The results from this work should be considered in longitudinal study designs and comparable prospective motion correction investigations are warranted in cases of marked head movement.

## 1. Introduction

Magnetic resonance imaging (MRI)-based neurochemical, volumetric, and morphological (e.g. cortical thickness, surface area) indices are often used to explore research questions in neuroscience. For example, existing large-scale MRI studies examining patients with schizophrenia, Alzheimer’s disease, major depressive disorder, and autism spec-

trum disorder have previously identified neurochemical, volumetric, and morphological abnormalities, as compared to healthy controls (Schmaal et al., 2016; van Erp et al., 2018, 2016; van Rooij et al., 2018). Outside of the context of neuropsychiatric disorders, longitudinal MRI made a significant impact on studies of brain maturation (Raznahan et al., 2014; Reardon et al., 2018) and ageing (Fjell et al., 2015; Tullo et al., 2019; Voineskos et al., 2015). As such, various MRI

\* Corresponding authors.

E-mail addresses: [plitmaneric@gmail.com](mailto:plitmaneric@gmail.com) (E. Plitman), [mallar@cobralab.ca](mailto:mallar@cobralab.ca) (M.M. Chakravarty).

<https://doi.org/10.1016/j.neuroimage.2021.118172>.

Received 16 June 2020; Received in revised form 7 May 2021; Accepted 8 May 2021

Available online 31 May 2021.

1053-8119/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

indices are being investigated to assess their potential role as biomarkers for neuropsychiatric disorders and their utility in clinical staging, prognostication, prediction of illness onset, and the development of normative trajectories, as well as to further the general understanding of neuroscience.

Compared to cross-sectional approaches, longitudinal MRI studies are often considered a superior strategy for biomarker investigations. Longitudinal brain imaging mitigates potential sources of confounding that may affect cross-sectional studies, such as participant heterogeneity and cohort effects. As a result, indices identified through longitudinal studies as relevant biomarkers in the pathophysiology of various neuropsychiatric disorders can be more readily assigned causality (Cannon et al., 2015; Chincarini et al., 2016; Jack et al., 2013; van Haren et al., 2011).

One concern that is pertinent to the completion of a longitudinal MRI study is the fact that MRI scanners often undergo hardware and software upgrades. It is conceivable that these upgrades may affect the measurement of various relevant MRI indices. When upgrades occur during the course of a longitudinal MRI investigation, there may be an impact on the compatibility of pre- and post-upgrade measures (Lee et al., 2019; Takao et al., 2013). Our research group recently underwent an upgrade for our 3T MRI scanner, from a Siemens MAGNETOM 3T Tim Trio system to a Siemens MAGNETOM 3T Prisma Fit system. This included hardware upgrades to the gradient, radiofrequency, and shimming systems as well as software and applications upgrades (Siemens, 2020).

Another important issue for longitudinal MRI studies (although relevant to all MRI studies) is subject motion. Subject motion during a three-dimensional MRI acquisition is well-known to cause artifacts (e.g. ghosting, blurring) by making k-space sampling inconsistent, which may be detrimental to the quality and utility of the scan (Tisdall et al., 2012). This is an important consideration for longitudinal work, as it is common practice amongst research groups to remove poor quality scans from further analysis. This represents a loss of significant resources, which may often influence the utility of other data collected from that participant over the course of a longitudinal study design. In addition, removing high-motion scans can represent a source of bias within the dataset and unresolved motion has been demonstrated to impact both volumetric and morphological estimates (Bedford et al., 2020; Reuter et al., 2015). The impact of subject motion may be mitigated through multiple approaches, which may be classified as prospective and retrospective. One prospective technique that has become increasingly common involves embedding volumetric navigators (vNavs) (i.e. fast-acquisition low-resolution 3D EPI images) within longer acquisition sequences (Tisdall et al., 2016, 2012). vNavs undergo rapid rigid registration to estimate changes in head position, effectively tracking and updating subject motion during an acquisition. Among the advantages of this approach is its efficiency, as the vNavs are distributed throughout existing sequences within “dead-times”, thus providing benefit at a negligible cost of time, contrast, and intensity. Also, vNav sequences may permit the re-acquisition of repetition times (TRs) within which a large quantity of motion exists. Notably, vNavs have been shown to lead to a reduction in morphometry variation attributed to motion (Tisdall et al., 2016). As a result, research groups within the field have altered their protocols to switch from standard structural sequences to ones with vNavs, which are currently also being employed in large-scale initiatives where participant motion is known to be problematic (Alexander-Bloch et al., 2016; Pardoe et al., 2016) such as the Adolescent Brain Cognitive Development study (Casey et al., 2018). Understanding the effects of switching to or adding a vNav sequence to a standard T1-weighted acquisition that is extensively used within the field (e.g. ADNI protocol), which may have differing parameters, is particularly relevant to longitudinal as well as multi-site work.

The goals of the current work were to: 1) investigate the impact of upgrading from a Siemens 3T Tim Trio system to a Siemens 3T MAGNETOM Prisma system on commonly used volumetric and morphometric structural imaging measures and proton magnetic resonance spec-

troscopy ( $^1\text{H-MRS}$ ) indices and 2) examine the reliability and similarity of volumetric and morphometric structural imaging measures in a sequence with vNavs alongside a standard T1-weighted acquisition sequence that closely aligns with the ADNI protocol (Bussy et al., 2021; Jack et al., 2008).

## 2. Methods

### 2.1. Participants

This study received approval by the Research Ethics Board of the Douglas Mental Health University Institute. Written informed consent was acquired from a total of nineteen individuals who completed the screening interview and were subsequently enrolled into the present study. Inclusion criteria were: capacity to consent; age 18-80; and fluency in English or French. Exclusion criteria included contraindications to MRI.

### 2.2. Study design

This study was conducted from June 2018 to September 2019, and consisted of two protocols (Fig. 1). The sample size within each investigation is detailed within Supplementary Table 1.

#### 2.2.1. Prisma Upgrade protocol

The first protocol was designed to assess the impact of the Prisma upgrade, and consisted of four MRI visits per subject: two using a Siemens MAGNETOM Tim Trio MRI scanner and two using the Siemens MAGNETOM Prisma MRI Fit scanner following the completion of an upgrade. To merit inclusion of a subject, the participant must have undergone at least one pre-upgrade and one post-upgrade acquisition. In this study protocol, fourteen healthy individuals participated, although one participant failed to complete any post-upgrade scans. Thus, data from thirteen individuals (mean age:  $28.62 \pm 6.14$  years, 6 females) were used for analysis. Nine subjects completed all four MRI visits, one subject completed two pre-upgrade visits and one post-upgrade visit, one subject completed one pre-upgrade visit and two post-upgrade visits, and two subjects completed one pre-upgrade visit and one post-upgrade visit. One participant who completed two pre-upgrade MPRAGE acquisitions only completed one pre-upgrade  $^1\text{H-MRS}$  acquisition.

#### 2.2.2. vNav protocol

The second study protocol consisted of two MRI visits and was principally designed to: 1) investigate vNav structural sequences alongside standard structural acquisitions and 2) acquire post-upgrade  $^1\text{H-MRS}$ . Due to the latter purpose, we aimed to include as many subjects as possible from the Prisma Upgrade protocol in the vNav protocol. Thirteen individuals (mean age:  $27.92 \pm 6.63$  years, 7 females) participated in the vNav protocol, eight of which also participated in the Prisma Upgrade protocol. MPRAGE acquisitions always occurred prior to vNav-MPRAGE acquisitions within a scanning session. Notably, vNav-MPRAGE acquisitions had different scanning protocol parameters (more detail in 2.3) compared to MPRAGE acquisitions; thus, these protocol changes as a whole make up the basis of our investigated comparison between vNav-MPRAGE and MPRAGE acquisitions. Finally, MPRAGE data acquired within the vNav protocol was used to supplement post-upgrade (i.e. Prisma-Prisma) analyses within the Prisma Upgrade protocol.

### 2.3. Magnetic resonance imaging acquisition

Participants were scanned at the Douglas Mental Health University Institute. In the Prisma Upgrade protocol, scanning procedures during MRI visits 1 and 2 were performed on a 3T Siemens Tim Trio MRI machine using a 32-channel head coil (software version: VB17A). For MRI visits 3 and 4, scanning procedures were performed on a 3T Siemens MAGNETOM Prisma MRI machine using a 32-channel

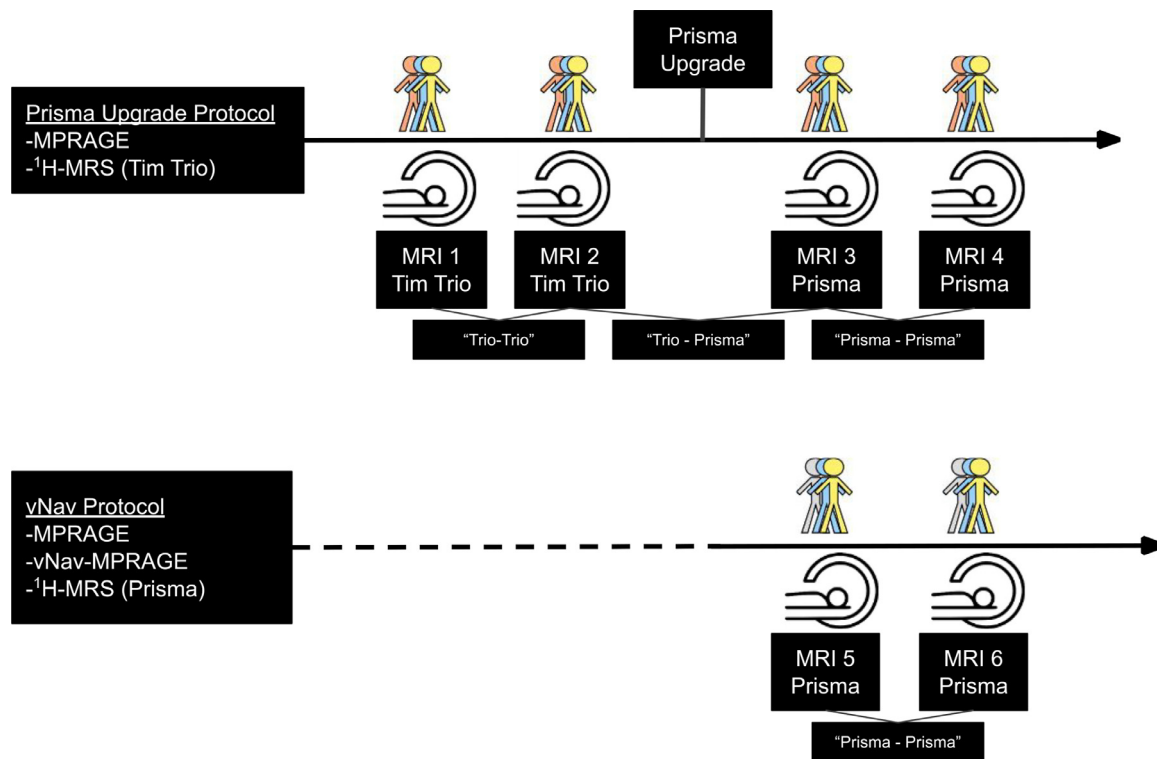


Fig. 1. Study design.

head coil (software version: VE11C). In each of these MRI visits, T1-weighted structural images were acquired (MPRAGE, TR=2300ms, TE=2.98ms (Tim Trio) and 2.01ms (Prisma), TI=900ms, flip angle=9°, bandwidth=238Hz/Px (Tim Trio) and 240Hz/Px (Prisma), voxel size=1.0 × 1.0 × 1.0mm, 192 × 240 × 256 matrix, GRAPPA = 2, prescan normalize). In the vNav protocol, all scanning procedures were performed on a 3T Siemens MAGNETOM Prisma MRI machine. During both of the MRI visits, MPRAGE (parameters noted above) and vNav-MPRAGE (TR=2300ms, TE=2.90 ms, TI=1070ms, flip angle=9°, bandwidth=240Hz/Px, voxel size=1.0 × 1.0 × 1.0mm, 176 × 240 × 256 matrix, GRAPPA = 2, prescan normalize) were acquired. vNav-MPRAGE sequences were provided by Massachusetts General Hospital.

#### 2.4. Magnetic resonance spectroscopy acquisition

In the Prisma Upgrade protocol, <sup>1</sup>H-MRS was acquired within the left dorsolateral prefrontal cortex (DLPFC) during MRI visits 1 and 2. As part of the vNav protocol, <sup>1</sup>H-MRS was acquired within the left DLPFC during both of the MRI visits. For <sup>1</sup>H-MRS scans in both the Prisma Upgrade protocol and the vNav protocol, the MPRAGE anatomical image was used to guide placement of a 2.5 × 2.5 × 1.5 cm<sup>3</sup> voxel in the left DLPFC. The voxel was first positioned such that when viewed in the sagittal plane, the centre of the MRS voxel in the A/P direction was aligned with the anterior extent of the genu of the corpus callosum. Next, when viewed in the coronal plane at the level of the genu of the corpus callosum, the voxel was positioned at the brain surface, approximately mid-way between the dorsomedial extent and the left-most lateral extent of the brain. To ensure maximal coverage of cortical grey matter in that region, the voxel was rotated about both the scanner y- and x-axes such that the largest face of the voxel (the 2.5 × 2.5 cm<sup>2</sup> face) was parallel to the surface of the brain. Short echo-time <sup>1</sup>H-MRS data were acquired using the SPECIAL sequence (Mekle et al., 2009) with TR=3500 ms (Tim Trio) and 3200 ms (Prisma), TE=8.5 ms, 4096 spectral points, 4000 Hz spectral width. 144 (Tim Trio) and 160 (Prisma) water-suppressed averages were acquired using the VAPOR water sup-

pression scheme (Tkáč et al., 1999), and an additional 8 non-water-suppressed averages were acquired for use in array coil reconstruction, lineshape correction, and optional signal referencing.

Prior to <sup>1</sup>H-MRS acquisition, localized shimming was performed on the volume of interest. In the Prisma Upgrade protocol, shimming was performed using the projection-based FASTESTMAP method (Gruetter and Tkáč, 2000), while in the vNav protocol, shimming was performed using the fieldmap-based GRE-SHIM method.

#### 2.5. Preprocessing

Unless otherwise specified, structural imaging analysis was done in the minc format on a high-performance computing cluster (SciNet). T1-weighted structural images were preprocessed using the minc-bpipe-library pipeline (<https://github.com/CobraLab/minc-bpipe-library>), which included a signal intensity correction (i.e. N4-correction (Tustison et al., 2010)), and procedures to exclude the neck and skull. Notably, gradient non-linearity correction was not applied on the scanner or in post-processing pipelines. <sup>1</sup>H-MRS analyses included the simulation of a site-specific basis set and data preprocessing performed using the FID-A toolkit (Simpson et al., 2017), the latter of which included removal of motion-corrupted averages, along with frequency and phase drift correction.

#### 2.6. Morphological analysis

Cortical thickness (CT) was estimated using the CIVET processing pipeline (Lerch and Evans, 2005) (version 2.1.0) on N4-corrected (Tustison et al., 2010) T1-weighted inputs (e.g. MPRAGE and vNav-MPRAGE). More details are included within the Supplementary Material. The mean CT was estimated in 31 bilateral regions of interest (ROIs), as defined by the Desikan-Killiany-Tourville atlas (Klein and Tourville, 2012), using unsmoothed “tlaplace” outputs. In addition, for complementary analyses, CT was estimated using the recon-all pipeline from FreeSurfer (version 6.0) (Dale et al.,

1999; Fischl et al., 1999) on native T1-weighted images (e.g. MPRAGE and vNav-MPRAGE) in the DICOM file format. FreeSurfer methodology and documentation has been previously detailed and is available at: <http://surfer.nmr.mgh.harvard.edu>. Parcellation of reconstructed surfaces from FreeSurfer rendered CT estimates for each participant in ROIs defined by the Desikan-Killiany-Tourville atlas (Klein and Tourville, 2012). For consistency in nomenclature with CIVET, lateral occipital, pars opercularis, pars orbitalis, and pars triangularis ROIs are referred to as the inferior occipital cortex, lateral frontal opercularis, lateral frontal orbitalis, and lateral frontal triangularis, respectively.

## 2.7. Volumetric analysis

Fully-automated segmentations of neuroanatomical structure volumes (SVs) (e.g. amygdala, globus pallidus, hippocampus, striatum, thalamus) were carried out using the Multiple Automatically Generated Templates (MAGeT-Brain) algorithm (Chakravarty et al., 2013; Pipitone et al., 2014). More details are included within the Supplementary Material. As above, FreeSurfer was used, wherein automated segmentation rendered cortical and subcortical volumes (Fischl et al., 2002). Given the complementary nature of this analysis, only volumes from structures investigated with MAGeT-Brain (e.g. amygdala, globus pallidus [i.e. pallidum], hippocampus, striatum [i.e. sum of caudate and putamen], thalamus [i.e. thalamus proper]), were used.

For CT and SV analyses, outputs were not excluded from analysis on the basis of quality.

## 2.8. Signal-to-Noise analysis

Signal-to-noise ratios (SNR) were estimated using noise\_estimate (minc-toolkit-v2 1.9.17) (Coupé et al., 2009) from native MPRAGE acquisitions in the MINC format that were included in the Prisma Upgrade protocol. noise\_estimate calculates the SNR value through a wavelet decomposition of the image to extract the object (signal) from Kmeans clustering of the lowest subband while the highest subband components are purely noise. The initial signal to noise estimate is then corrected for Rician noise distribution by iteratively solving an analytic expression relating signal to noise in MR until convergence. This method was originally developed for non-accelerated scans with stationary noise distributions, whereas the protocols used in this study are GRAPPA accelerated; however, coil count and acceleration remain constant in the study, so estimated SNR values, while not necessarily true values, are comparable.

## 2.9. <sup>1</sup>H-MRS analysis

Glutamate (Glu), glutamine (Gln), Glu + Gln (Glx), myo-inositol (Ins), total choline (Cho, sum of glycerophosphocholine and phosphocholine), N-acetylaspartate (NAA), N-acetylaspartylglutamate (NAAG), total N-acetylaspartate (tNAA, sum of NAA and NAAG), gamma-Aminobutyric acid (GABA), glutathione (GSH), and lactate (Lac) were estimated from preprocessed <sup>1</sup>H-MRS data using LCModel version 6.3-1H (Provencher, 2001). Neurometabolite levels were referenced to total creatine (sum of creatine and phosphocreatine) levels. A basis set of neurometabolites, obtained via simulation in FID-A (Simpson et al., 2017) according to the *in vivo* scan parameters, was used. A custom basis set that was used is described within the Supplementary Material. In the current work, no cutoffs were employed for %SD, full-width at half maximum (FWHM), or SNR. Notably, FWHM and SNR were calculated using FID-A (Simpson et al., 2017). FWHM represents the FWHM of the water peak in the water unsuppressed scan and SNR is the signal of the NAA peak in the water suppressed scan divided by the standard deviation of the noise between 0 and 2 ppm in the same spectrum.

## 2.10. Statistical analyses

Analyses were performed in R version 3.5.0 and can be further divided into “test-retest reliability analyses” and “similarity analyses”; the former sought to examine the test-retest reliability across two sessions using the same imaging sequence, whereas the latter sought to investigate the similarity between two different imaging sequences within the same study visit. Intraclass correlation coefficients (ICCs) were determined to test reliability using the ICC function in the *psych* 1.9.12.31 package. ICCs (3,1) (i.e. consistency) for which  $p < 0.1$  were interpreted according to previously established criteria (Cicchetti, 1994):

- “excellent”: 1.00-0.75
- “good”: 0.74-0.60
- “fair”: 0.59-0.40
- “poor”: 0.39-0.00

ICCs (2,1) (i.e. absolute agreement) are reported in Tables and the Supplementary Material but were not used for interpretation. Also, absolute percentage difference (APD) was calculated according to the following formula:  $[(\text{timepoint 1 value} - \text{timepoint 2 value}) / ((\text{timepoint 1 value} + \text{timepoint 2 value}) / 2)] \times 100$ . Notably, APDs describe the change in a variable of interest without the assessment of any confounding variable. Finally, linear mixed-effects model analyses used the *lme4* 1.1-23 and *lmerTest* 3.0-1 packages to test whether the difference between two timepoints or sequences was statistically significant, while confounding for variability in subjects by using subject as a random intercept. Here, a Bonferroni correction for multiple comparisons was employed; thus, a significance level of  $p < 0.00081$  ( $0.05 / 62$  ROIs) was employed for CT linear mixed-effects analyses, whereas a significance level of  $p < 0.005$  ( $0.05 / 10$ ) was employed for both SV and <sup>1</sup>H-MRS linear mixed-effects analyses. More detail regarding each study protocol is provided below.

### 2.10.1. Test-retest reliability analyses

For the Prisma Upgrade protocol, ICCs and APDs were calculated for the following pairs: visit 1 and visit 2 (i.e. Tim Trio test-retest reliability, henceforth “Trio-Trio”), visit 2 and visit 3 (i.e. impact of Prisma upgrade, henceforth “Trio-Prisma”), and visit 3 and visit 4 (i.e. Prisma test-retest reliability, henceforth “Prisma-Prisma”). For the vNav protocol, ICCs were calculated between visit 5 and 6 (i.e. Prisma-Prisma). Of note, Trio-Prisma pairs were selected according to the criterion of proximity in time. The breakdown of data within each of the pairs noted above and the duration between acquisitions is in Supplementary Table 1. It deserves mention that MPRAGE reliabilities for CT and SV were recalculated for the first and second visits of the vNav protocol to serve as comparison for vNav reliabilities, using the same participant pool to do so. In addition, linear mixed-effects models were used to test differences between acquisitions in the pairs noted above, with visit as a fixed effect and subject as a random effect (*Measure of Interest* ~ *Visit* + (1|*Subject*)).

### 2.10.2. Similarity analyses

As part of the primary aim of the vNav protocol, similarity analyses compared standard structural sequences with vNav sequences within the same MRI session. Thus, similarity analyses tested the following pair: MPRAGE and vNav-MPRAGE. The breakdown of data within each pair is further displayed in Supplementary Table 1. For each participant, pairs were preferentially selected from the first visit of the vNav protocol. As above, ICCs and APDs were calculated for each pair, and linear mixed-effects models were used to test differences between acquisitions, with acquisition as a fixed effect and subject as a random effect (*Measure of Interest* ~ *Acquisition* + (1|*Subject*)).

### 2.10.3. Signal-to-Noise analyses

Using only the data from Trio-Prisma pairs noted above, a linear mixed-effects model was used to compare SNR between Tim Trio and Prisma acquisitions, with scanner as a fixed effect and subject as a random effect.



**Table 1**  
Trio-Prisma reliability measures for structure volume estimated by the MAGeT-Brain pipeline.

ROI	MAGeT-Brain Left Hemisphere					Right Hemisphere				
	APD	ICC (3,1)	p-value (ICC)	t-value (lmer)	p-value (lmer)	APD	ICC (3,1)	p-value (ICC)	t-value (lmer)	p-value (lmer)
Amygdala	2.53	0.98	<0.001	-4.05	0.002*	1.62	0.98	<0.001	-1.36	0.199
Hippocampus	2.67	0.96	<0.001	-3.61	0.004*	2.67	0.95	<0.001	-1.94	0.077
Globus Pallidus	3.01	0.98	<0.001	-5.70	<0.001*	3.11	0.95	<0.001	-3.21	0.008
Striatum	0.65	1.00	<0.001	2.45	0.030	0.56	1.00	<0.001	-0.88	0.396
Thalamus	3.61	0.97	<0.001	7.32	<0.001*	3.76	0.98	<0.001	7.91	<0.001*

APD, absolute percentage difference; ICC, intraclass correlation coefficient; lmer, linear mixed-effects model; ROI, region of interest. Note: positive t-values represent structure volume increases and presented p-values are uncorrected. Asterisk indicates significance after Bonferroni correction.

### 2.11. Power analyses

In order to assess the effects of a scanner upgrade in a real-world scenario, power analyses were conducted on simulated experiments. Selecting right hippocampus, precuneus, inferior occipital, rostral middle frontal, and superior temporal ROIs, we used previously estimated volume and CT from participants scanned either in Prisma-Prisma or Trio-Trio conditions as healthy controls. We duplicated the dataset and artificially created a “patient” group with a 5% increase in volume or CT at the second timepoint. Next, additional subjects were simulated to estimate the sample size necessary to detect the real 5% increase, given a scanner upgrade. These power calculations were created using the *SIMR* package/1.0.4, and assessed the interaction between group and timepoint as the fixed effect of interest. *SIMR* power calculations are based on Monte Carlo simulations, allowing users to simulate based on linear mixed effects modules with the *lme4* packages (Green and MacLeod, 2016). The full model included the interaction as the fixed effect and subject and scanner as random effects.

## 3. Results

### 3.1. Prisma Upgrade protocol

#### 3.1.1. Cortical thickness

ICCs, APDs, and linear mixed-effects analyses results reflecting test-retest reliability for CT measures estimated from CIVET and FreeSurfer across Trio-Trio, Trio-Prisma, and Prisma-Prisma investigations are shown in Supplementary Tables 2 - 7. In the aforementioned analyses, reliability was found to be “excellent” or “good”, with the exception of those detailed below. For Trio-Trio investigations, reliability was “fair” within the right lateral orbitofrontal (CIVET) ROI. For Trio-Prisma investigations, reliability was “fair” within right lingual gyrus (CIVET), left transverse temporal (FreeSurfer), right rostral anterior cingulate (FreeSurfer), and right inferior occipital (FreeSurfer) ROIs, while reliability was “poor” within the right isthmus cingulate gyrus (CIVET) ROI. ICC results using CIVET and FreeSurfer are visualized in Fig. 2 and Supplementary Figure 1, respectively.

Results of linear mixed-effects analyses demonstrated differences between Tim Trio and Prisma estimates. Using CIVET, a significant effect of the upgrade was found for the following ROIs: bilateral lateral orbitofrontal, bilateral medial orbitofrontal, right inferior temporal, right insula, right lateral frontal orbitalis, right middle temporal, and right superior temporal (Fig. 3). Using FreeSurfer, a significant effect of the upgrade was found for the following ROIs: bilateral pericalcarine, left caudal anterior cingulate, left superior parietal, and right insula (Supplementary Figure 2). Linear mixed-effect analyses results using CIVET and FreeSurfer are illustrated in Fig. 4 and Supplementary Fig. 3, respectively.

#### 3.1.2. Structure volume

ICCs, APDs, and linear mixed-effects analyses results reflecting test-retest reliability for SV measures estimated from MAGeT-Brain and

**Table 2**

Trio-Prisma reliability measures for <sup>1</sup>H-MRS indices.

Measure	APD	ICC (3,1)	p-value (ICC)	t-value (lmer)	p-value (lmer)
Cho	7.92	0.73	0.013	2.59	0.036
GABA	15.40	0.30	0.220	0.46	0.659
Gln	13.60	0.40	0.144	-1.82	0.112
Glu	2.82	0.90	<0.001	0.67	0.523
Glx	3.08	0.90	<0.001	-0.64	0.541
GSH	4.98	0.61	0.040	0.24	0.820
Ins	10.10	0.62	0.037	3.87	0.006
Lac	19.20	0.32	0.203	-0.14	0.893
NAA	5.17	0.77	0.008	-2.69	0.031
NAAG	27.60	0.33	0.195	1.91	0.098
tNAA	3.02	0.81	0.004	-1.83	0.109
FWHM	13.40	0.21	0.292	-1.60	0.153
SNR	12.60	0.36	0.168	-0.08	0.936

APD, absolute percentage difference; ICC, intraclass correlation coefficient; lmer, linear mixed-effects model. Note: positive t-values represent increases and presented p-values are uncorrected.

FreeSurfer across Trio-Trio, Trio-Prisma, and Prisma-Prisma investigations are shown in Table 1 (MAGeT-Brain and Trio-Prisma) and Supplementary Tables 8 - 13. All MAGeT-Brain analyses rendered “excellent” reliability.

Using FreeSurfer, reliability was found to be “excellent” or “good”, with the exception of those detailed below. Reliability with FreeSurfer across Trio-Prisma investigations was “fair” within the left amygdala and right thalamus, and did not reach statistical significance within the bilateral globus pallidus, left hippocampus, and left thalamus. Further, reliability with FreeSurfer across Prisma-Prisma investigations was “fair” within the left globus pallidus and “poor” within the right globus pallidus.

Results of linear mixed-effects analyses demonstrated differences between Tim Trio and Prisma estimates. Using MAGeT-Brain, a significant effect of the upgrade was found for the following ROIs: bilateral thalamus, left amygdala, left globus pallidus, and left hippocampus (Fig. 5). Using FreeSurfer, a significant effect of the upgrade was found for the following ROIs: bilateral amygdala, bilateral globus pallidus, and bilateral striatum (Supplementary Figure 4).

#### 3.1.3. Proton magnetic resonance spectroscopy

ICCs, APDs, and linear mixed-effects analyses results reflecting test-retest reliability for <sup>1</sup>H-MRS measures across Trio-Trio, Trio-Prisma, and Prisma-Prisma investigations are shown in Table 2 (Trio-Prisma) and Supplementary Tables 14 - 16. All <sup>1</sup>H-MRS metrics achieved “excellent” or “good” reliability with the exception of those detailed below. For Trio-Trio investigations, reliability was “fair” for Lac and SNR, and did not reach statistical significance for GABA, NAAG, and FWHM. For Trio-Prisma investigations, reliability did not reach statistical significance for GABA, Gln, Lac, NAAG, FWHM, and SNR. For Prisma-Prisma investigations, reliability was “fair” for GABA, NAAG and SNR, and did not reach statistical significance for GSH and Lac. Results of linear-mixed effects analyses were not statistically significant.

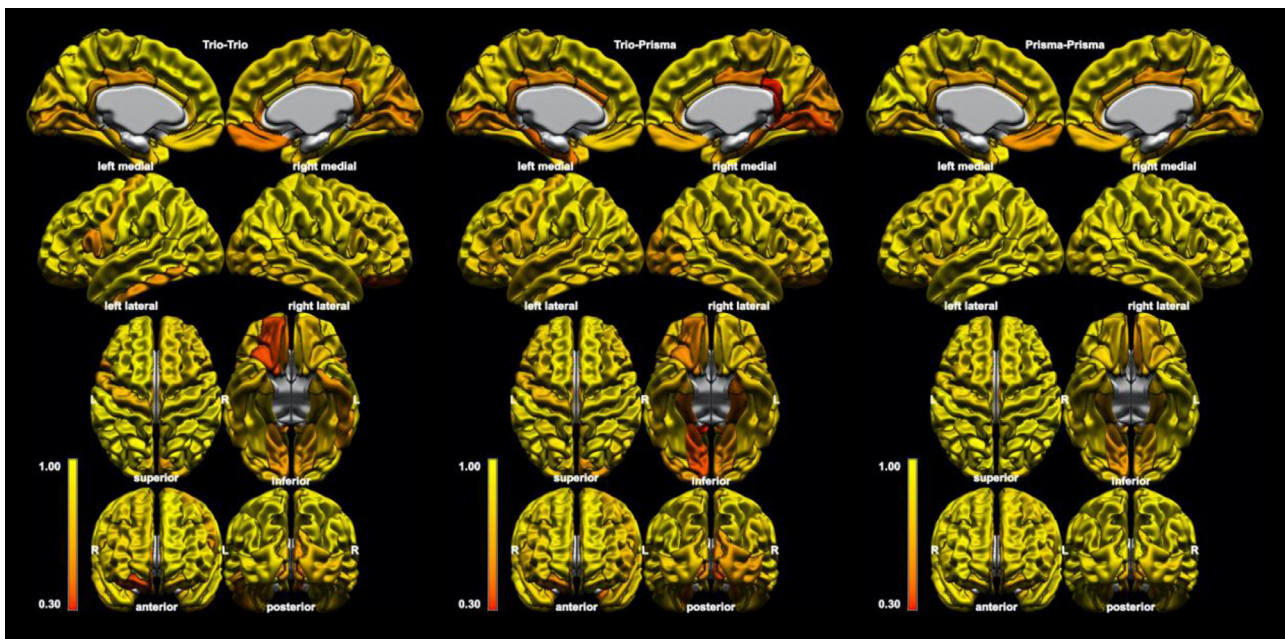


Fig. 2. Intra-class correlation coefficients of cortical thickness estimated by the CIVET pipeline within regions of interest defined by the Desikan-Killiany-Tourville atlas (black boundaries). Results are shown for pairs consisting of two T1-weighted images from a Tim Trio machine (left), Prisma machine (right), and one of each (middle).

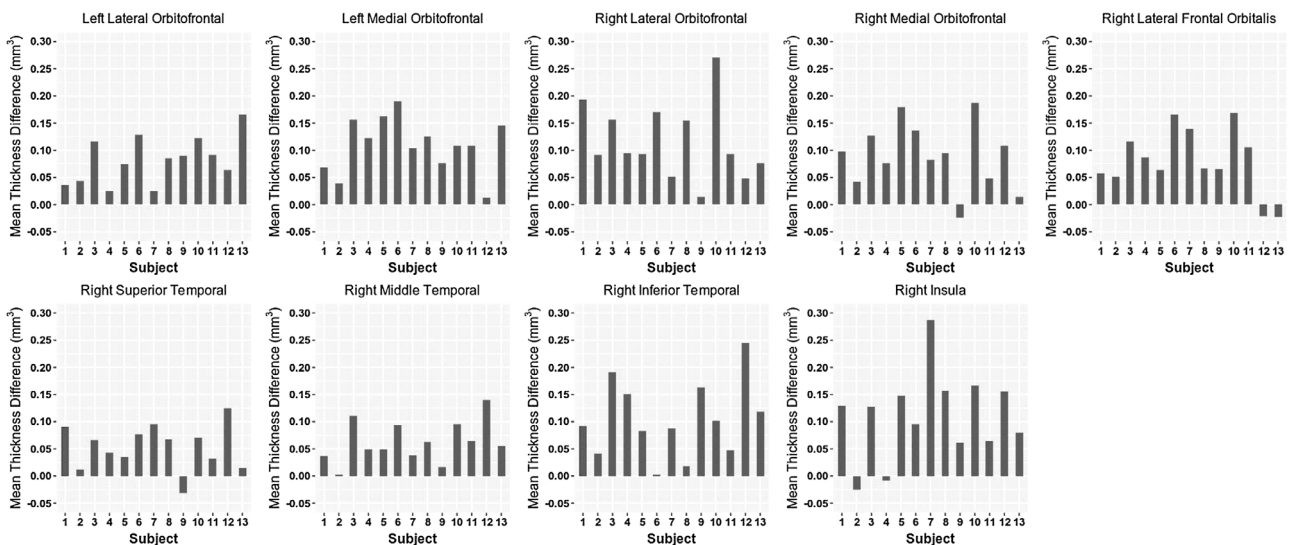


Fig. 3. Cortical thickness estimated by the CIVET pipeline within regions of interest that achieved statistical significance in linear mixed-effects analyses. Mean thickness difference represents Prisma estimate minus Tim Trio estimate.

Table 3

MPRAGE and vNav-MPRAGE similarity measures for structure volume estimated by the MAgE-T-Brain pipeline.

ROI	MAGeT-Brain Left Hemisphere				Right Hemisphere					
	APD		p-value (ICC)	t-value (lmer)	APD	ICC (3,1)		p-value (ICC)	t-value (lmer)	p-value (lmer)
	APD	ICC (3,1)				ICC (3,1)	ICC (3,1)			
Amygdala	2.02	0.98	<0.001	-0.73	0.477	1.87	0.98	<0.001	-0.24	0.813
Hippocampus	2.29	0.97	<0.001	1.24	0.238	2.51	0.98	<0.001	1.45	0.172
Globus Pallidus	1.52	0.98	<0.001	3.10	0.009	1.60	0.97	<0.001	2.36	0.036
Striatum	0.44	1.00	<0.001	-0.52	0.610	0.28	1.00	<0.001	1.28	0.224
Thalamus	0.90	0.98	<0.001	-0.59	0.568	0.73	0.99	<0.001	-0.97	0.353

APD, absolute percentage difference; ICC, intraclass correlation coefficient; lmer, linear mixed-effects model; ROI, region of interest. Note: positive t-values represent structure volume increases in vNav-MPRAGE and presented p-values are uncorrected.

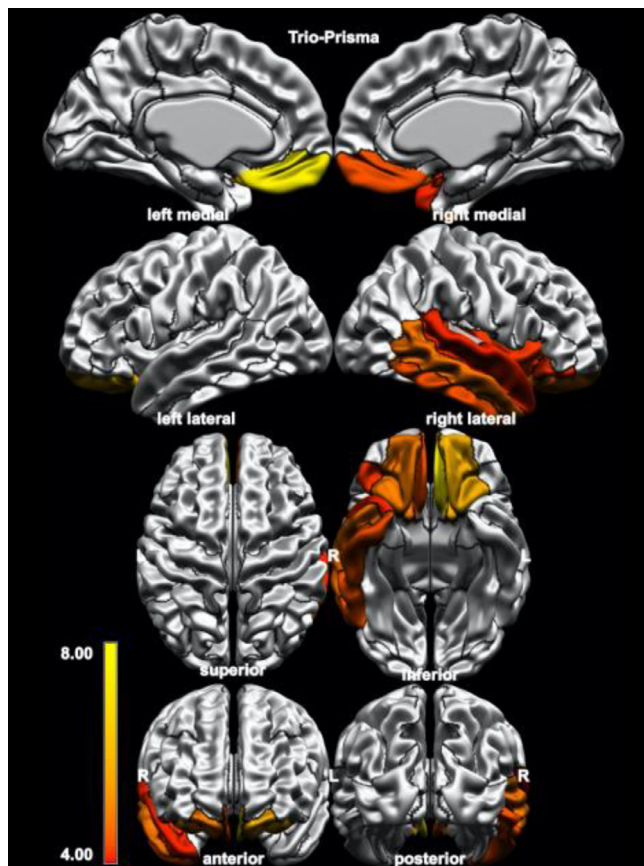


Fig. 4. Cortical thickness estimated by the CIVET pipeline within regions of interest that achieved statistical significance in linear mixed-effects analyses. Visualization represents t-statistics from linear mixed-effect analyses assessing differences between Tim Trio and Prisma acquisitions.

### 3.2. vNav protocol

#### 3.2.1. Cortical thickness

**3.2.1.1. Test-retest reliability.** ICCs, APDs, and linear mixed-effects analyses results reflecting test-retest reliability for CT measures estimated from CIVET and FreeSurfer in MPRAGE and vNav-MPRAGE sequences are shown in Supplementary Tables 17–20. Using CIVET, reliability was found to be “excellent” or “good”, with the exception of those detailed below. For MPRAGE investigations, reliability was “fair” within the left lateral orbitofrontal ROI. For vNav-MPRAGE investigations, reliability was “fair” within the left lateral frontal orbitalis, right isthmus cingulate, right pericalcarine, and right posterior cingulate ROIs. FreeSurfer analyses rendered “excellent” or “good” reliability, with the exception of the right rostral anterior cingulate ROI in vNav-MPRAGE sequences, which was found to be “fair”. ICC results using CIVET and FreeSurfer are illustrated in Fig. 6 and Supplementary Figure 5, respectively. Results of linear-mixed effects analyses were not statistically significant.

**3.2.1.2. Similarity analyses.** ICCs, APDs, and linear mixed-effects analyses results reflecting similarity of CT measures acquired from MPRAGE and vNav-MPRAGE sequences, as estimated by CIVET and FreeSurfer, are shown in Supplementary Tables 21 and 22, respectively. In the aforementioned analyses, reliability was found to be “excellent” or “good”, with the exception of those detailed below. Reliability was “fair” within the right entorhinal cortex (CIVET), “poor” within the right isthmus cingulate (CIVET), and did not reach statistical significance within the right entorhinal cortex (FreeSurfer). ICC results using CIVET and FreeSurfer are illustrated in Fig. 6 and Supplementary Fig. 5, respectively.

Results of linear mixed-effects analyses rendered differences between CT estimates from MPRAGE and vNav-MPRAGE sequences. Using CIVET, a significant effect of sequence type was found for the following ROIs: bilateral parahippocampal, left medial orbitofrontal, left postcentral, right inferior parietal, right middle temporal, right precentral, right precuneus, and right superior temporal (Fig. 7). Using FreeSurfer, a significant effect of sequence type was found for the following ROIs: bilateral postcentral, left inferior parietal, left superior parietal, left transverse temporal, right inferior occipital, right lingual, right lateral frontal triangularis, right pericalcarine, right rostral middle frontal, and right superior frontal (Supplementary Figure 6). Linear mixed-effect analyses results using CIVET and FreeSurfer are further illustrated in Fig. 8 and Supplementary Fig. 7, respectively.

#### 3.2.2. Structure volume

**3.2.2.1. Test-retest reliability.** ICCs, APDs, and linear mixed-effects analyses results reflecting test-retest reliability for SV measures estimated from MAGeT-Brain and FreeSurfer in MPRAGE and vNav-MPRAGE sequences are shown in Supplementary Tables 23–26. All analyses rendered “excellent” or “good” reliability. Using MAGeT-Brain, results of linear-mixed effects analyses were not statistically significant. Using FreeSurfer, a significant effect of visit was found in the right amygdala for vNav-MPRAGE sequences.

**3.2.2.2. Similarity analyses.** ICCs, APDs, and linear mixed-effects analyses results reflecting similarity of SV measures acquired from MPRAGE and vNav-MPRAGE sequences, as estimated by MAGeT-Brain and FreeSurfer, are shown in Supplementary Tables 27 and 28, respectively. All analyses rendered “excellent” reliability. Using MAGeT-Brain, results of linear-mixed effects analyses were not statistically significant. Using FreeSurfer, a significant effect of sequence was found in the left hippocampus and right thalamus (Supplementary Fig. 8).

### 3.3. Signal-to-Noise analyses

A significant increase in SNR was identified following the scanner upgrade (Tim Trio:  $20.07 \pm 0.66$ , Prisma:  $24.30 \pm 1.32$ ,  $t = 13.98$ ,  $P_{\text{uncorrected}} < 0.001$ ) (Fig. 9).

### 3.4. Power analyses

Simulations revealed that, for all investigated structures, depending on the structure and pipeline used for estimation, 25–35 participants are sufficient to achieve 80% power at estimating a 5% increase in volume/CT comparing the simulated group to the acquired data (Supplementary Figs. 9 and 10).

## 4. Discussion

The present study aimed to: 1) investigate the impact of an upgrade from a Siemens Tim Trio to a Siemens Prisma MRI scanner and 2) compare a sequence with vNavs against a standard T1-weighted acquisition sequence that is extensively used within the field. While the results contained in this manuscript do not generalize to different vendors, the methods used provide a template for future investigations. To the best of our knowledge, the current work represents the largest sample used to assess the impact of the Prisma upgrade on volumetric and morphological indices, the first to use CIVET and MAGeT-Brain to investigate this research question, and the first to examine neurochemical indices. Our results suggest high reliability in most investigated MRI outputs across the Prisma upgrade, although suboptimal reliability was observed in a minority of ROIs. Further, in certain ROIs, despite high reliability, CT and SV estimates differed across the Prisma upgrade. In addition, the present study is the first to explore vNav reliability and similarity analyses across multiple MRI visits, and use CIVET and MAGeT-Brain to address this research question, albeit in a sample that may not be



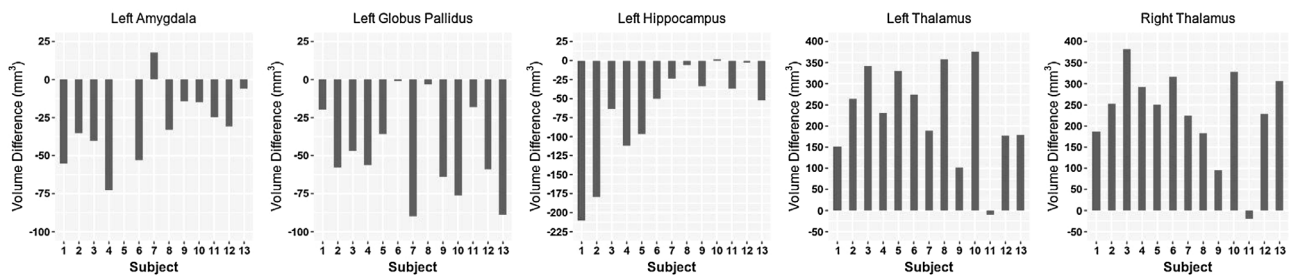


Fig. 5. Structure volume estimated by the MAGE-T-Brain pipeline within regions of interest that achieved statistical significance in linear mixed-effects analyses. Volume difference represents Prisma estimate minus Tim Trio estimate.

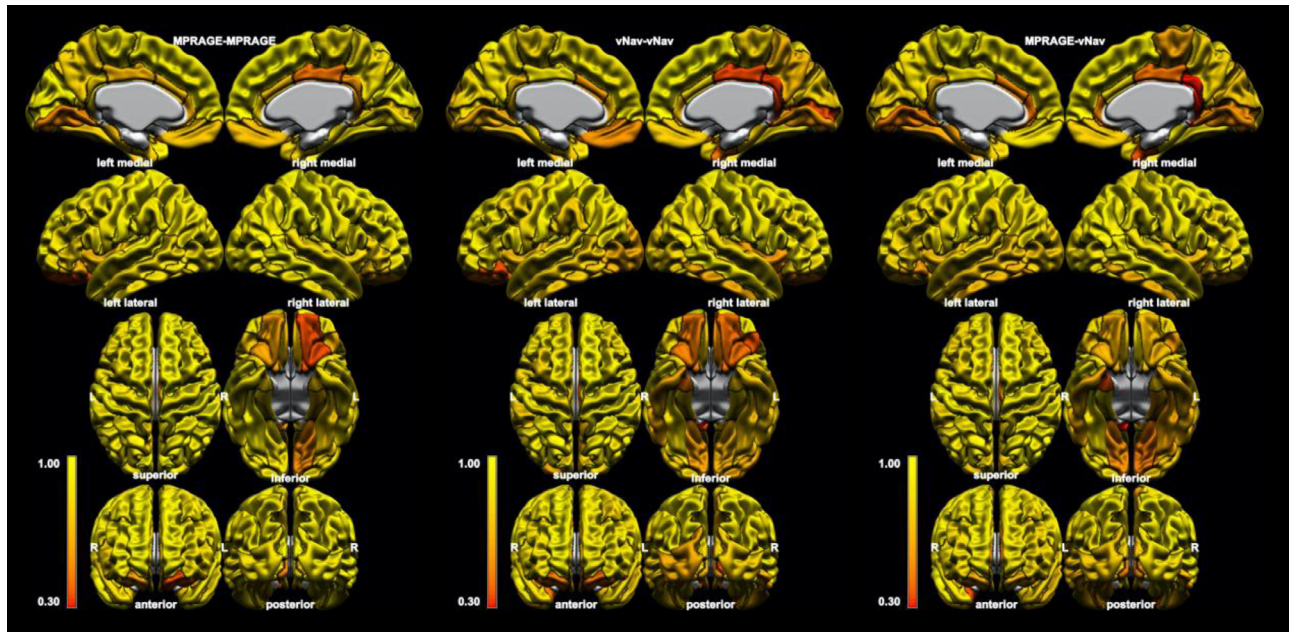


Fig. 6. Intra-class correlation coefficients of cortical thickness estimated by the CIVET pipeline within regions of interest defined by the Desikan-Killiany-Tourville atlas (black boundaries). Results are shown for pairs consisting of two T1-weighted images (left), two volumetric navigator T1-weighted images (middle), and one of each (right).

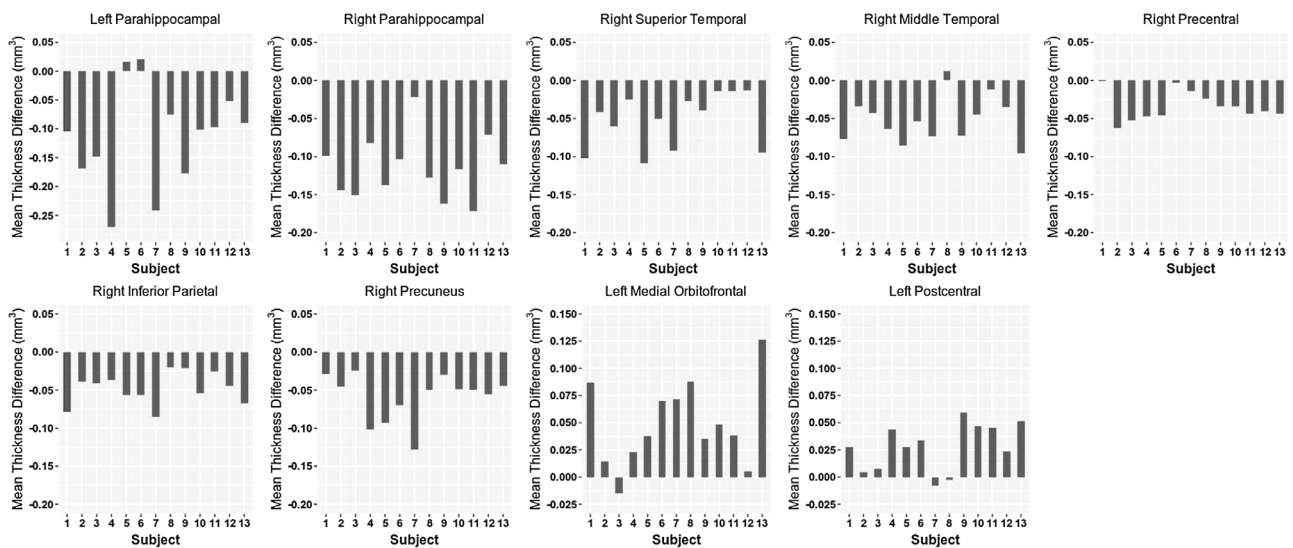
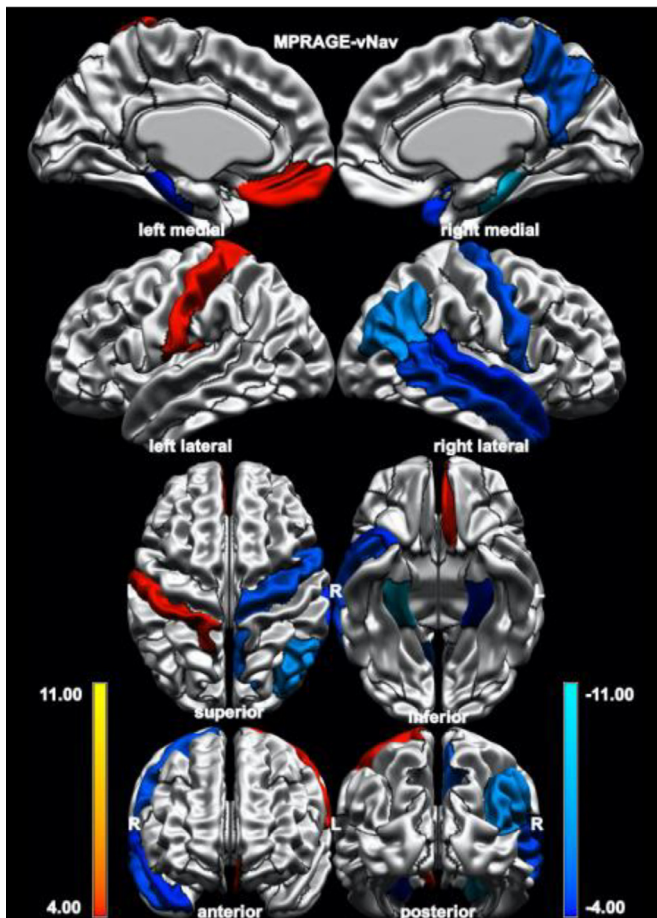


Fig. 7. Cortical thickness estimated by the CIVET pipeline within regions of interest that achieved statistical significance in linear mixed-effects analyses. Mean thickness difference represents vNav-MPRAGE estimate minus MPRAGE estimate.

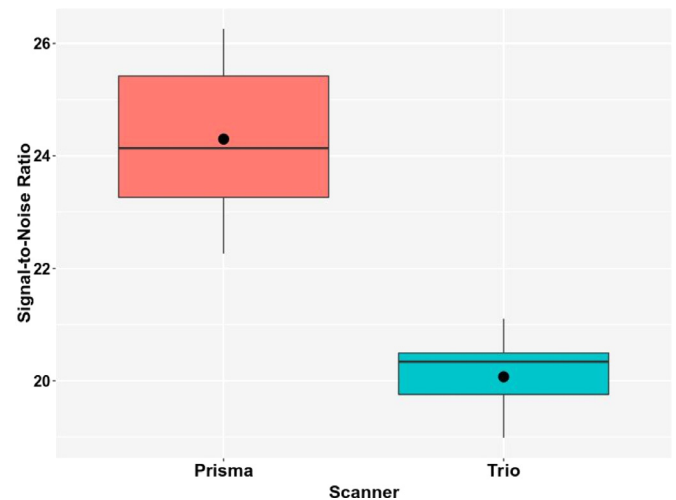




**Fig. 8.** Cortical thickness estimated by the CIVET pipeline within regions of interest that achieved statistical significance in linear mixed-effects analyses. Visualization represents t-statistics from linear mixed-effect analyses assessing differences between MPRAGE and vNav-MPRAGE acquisitions.

representative of the clinical population in whom prospective motion correction would be most beneficial. The current work identified certain ROIs within which MPRAGE and vNav-MPRAGE sequences were most dissimilar, although it is noteworthy that differing contrast parameters may also be of influence.

The identification of particular ROIs within which the Prisma upgrade impacted reliability is relevant towards longitudinal MRI investigations. In the present work, we observed “poor” reliability in right isthmus cingulate gyrus CT and “fair” reliability in right lingual gyrus CT, as estimated by CIVET, and “fair” reliability of CT within the left transverse temporal, right rostral anterior cingulate, and right inferior occipital ROIs, as estimated by FreeSurfer. Also, SV acquired using FreeSurfer in the left amygdala and right thalamus rendered “fair” reliability, and ICCs within other structures did not reach statistical significance. For future longitudinal MRI investigations, interpretation of estimates of CT and SV within these ROIs merits caution, as they may be impacted in an inconsistent manner by the Prisma upgrade. Notably, it is possible that low reliability observed for CT and SV estimates across the Prisma upgrade in these ROIs may be in part attributable to substandard reliability of these measures in general. To this effect, the examination of Trio-Trio and Prisma-Prisma reliability complements the interpretation of the Trio-Prisma results. For example, in Trio-Trio investigations, CIVET-estimated right lateral orbitofrontal CT had “fair” reliability, and in Prisma-Prisma investigations, FreeSurfer-estimated bilateral globus pallidus volume had “fair” or “poor” reliability. Interestingly, cortical ROIs with lower ICCs include those believed to be affected by distur-



**Fig. 9.** Signal-to-Noise Ratios in Siemens Prisma and Tim Trio scanners. Center circle reflects mean value.

tions (e.g. orbitofrontal areas) as well as those with a thin cortex that may contribute to poorer segmentation (e.g. V1). Thus, low reliability observed in these regions within Trio-Prisma pairs may not solely be a result of the upgrade.

Despite demonstrating “excellent” or “good” reliability across the Prisma upgrade, CT increases, as estimated by CIVET, were identified in 5 frontal ROIs (i.e. bilateral lateral orbitofrontal, bilateral medial orbitofrontal, right lateral frontal orbitalis), 3 right temporal ROIs (i.e. inferior temporal, middle temporal, superior temporal), and the right insula. Likewise, CT increases, as estimated by FreeSurfer, were identified within the left caudal anterior cingulate and the right insula; however, CT decreases were identified using FreeSurfer within 3 parietal ROIs (i.e. bilateral pericalcarine, left superior parietal). CT increases within frontal brain regions such as the medial orbitofrontal and lateral orbitofrontal ROIs were also observed using FreeSurfer, yet failed to achieve statistical significance after correction for multiple comparisons. Furthermore, the frontal and temporal CT increases identified by CIVET, along with the parietal CT decreases and temporal and cingulate CT increases identified by FreeSurfer, may be related to SNR increases that resulted from the Tim Trio to Prisma upgrade. This phenomenon is in line with previous studies that have reported upon volume increases attributed to SNR increases following MRI scanner upgrades (Lee et al., 2019; Potvin et al., 2019; Shuter et al., 2008). Notably, orbitofrontal regions in particular are influenced by magnetic susceptibility differences in the air-tissue interface that exists between the nasal sinus cavity and the lower frontal cortex; this phenomenon similarly applies to the temporal cortices and is supported by magnetic field  $B_0$  mapping (de Graaf, 2019). It deserves mention that the influence of protocol changes (e.g. TE reduction) and non-SNR factors (e.g. gradient non-linearity (Jovicich et al., 2006)) related to the upgrade on the identified structural alterations cannot be discounted in the current work.

The findings of increased CT within frontal and cingulate ROIs and decreased CT within parietal ROIs following the Prisma upgrade is consistent with current literature. In particular, Potvin and colleagues similarly investigated an upgrade from a Siemens Trio to the Prisma on three healthy male volunteers at three separate sites (Potvin et al., 2019). The authors analyzed T1-weighted images using FreeSurfer and, akin to the present study, parcellated ROIs using the DKT atlas. Increased thickness estimates were observed in several ROIs, predominantly within frontal regions, in addition to cingulate ROIs, and decreased thickness estimates were observed in left superior parietal and left entorhinal ROIs. Likewise, Takao and colleagues used voxel-based morphometry to identify gray matter volume alterations (both increases and decreases) following a GE scanner upgrade from HDx to HDxt, which were identified to

be comparable to the effects seen between two different scanners that were the exact same model (Takao et al., 2013). In contrast, Han and colleagues scanned five participants before and after an upgrade from a Siemens Magnetom Sonata to a Magnetom Avanto, noting that thickness measurements, as assessed by FreeSurfer, were largely unaffected (Han et al., 2006); however, in a separate protocol, the authors did identify CT increases with increasing field strength.

In the present study, as assessed by MAGeT-Brain, decreases in left amygdala, left globus pallidus, and left hippocampus volumes, along with increases in bilateral thalamus volumes, were identified across the Prisma upgrade. Similarly, decreases in bilateral amygdala, bilateral globus pallidus, and bilateral striatum volumes were identified across the Prisma upgrade using FreeSurfer. These findings merit comparison to those of Potvin and colleagues, who identified increases in right amygdala, right caudate, left pallidum, and right putamen volumes (Potvin et al., 2019). Interestingly, the authors identified a negative correlation between subcortical contrast-to-noise ratio and subcortical volume. Jovicich and colleagues reported that subcortical, ventricular, and intracranial volumetric measurements, as assessed by FreeSurfer, were minimally influenced by scanner upgrades (Siemens Sonata to Siemens Avanto and Siemens Trio to Siemens TrioTIM) and acquisition sequence (MPRAGE compared to Multi-echo FLASH), which was contrasted by variations in field strengths (1.5T compared to 3T) that did render a volume difference bias (Jovicich et al., 2009).

In the vNav protocol of the present study, the comparison of vNav-MPRAGE and MPRAGE sequences was not designed to assess the effectiveness of vNav-MPRAGE sequences as a motion correction technique, but rather to examine whether their use should be expected to impact estimates of volumetric and morphological indices, in a study with the expectation of a lack of substantial participant motion. Overall, the current work suggests high reliability (i.e. similarity) in most volumetric and morphological indices between vNav-MPRAGE and MPRAGE sequences. However, similarity analyses comparing vNav-MPRAGE and MPRAGE sequences demonstrated reduced reliability in CT estimates within the right entorhinal cortex (CIVET and FreeSurfer) and the right isthmus cingulate (CIVET). Notably, the finding of low reliability within the right cingulate ROIs between these sequences may be in part explained by the results of test-retest reliability analyses of the vNav-MPRAGE sequences, which similarly identified low reliability within cingulate areas. This may be attributed to poor midline classification, which may be a consequence of an asymmetric model not having been used for CT analyses. Along with the inclusion of vNavs, the influence of differing contrast parameters between vNav-MPRAGE and MPRAGE sequences (representing common modifications within the field), should be wholly considered and findings should be contextualized with those from work that matched MPRAGE and vNav-MPRAGE protocols parameter-for-parameter (Tisdall et al., 2012). In the current study, the ICCs between MPRAGE and vNav-MPRAGE were larger in certain ROIs than in MPRAGE-MPRAGE and vNav-MPRAGE-vNav-MPRAGE pairs. One potential explanation for this is that the former were acquired within the same scanning session, within a few minutes of one another. In contrast, acquisitions within MPRAGE-MPRAGE and vNav-MPRAGE-vNav-MPRAGE pairs were acquired within different scanning sessions. Collectively, these findings are applicable to studies that are composed of one participant subgroup that is more likely to move during an MRI scan and one participant subgroup that is more likely to remain still (e.g. case-control studies comparing patients with autism spectrum disorder and healthy controls).

CT estimated using CIVET was lower in vNav-MPRAGE compared to MPRAGE sequences within 4 temporal ROIs (i.e. bilateral parahippocampal, right superior temporal, right middle temporal), and right precentral, right inferior parietal, and right precuneus ROIs, yet was higher in vNav-MPRAGE compared to MPRAGE sequences within left medial orbitofrontal and left postcentral ROIs. In contrast, CT estimated using FreeSurfer was higher in vNav-MPRAGE compared to MPRAGE sequences within 5 parietal ROIs (i.e. bilateral postcentral, left superior

parietal, left inferior parietal, right pericalcarine), 3 frontal ROIs (i.e. right superior frontal, right rostral middle frontal, right lateral frontal triangularis), and left transverse temporal and right lingual ROIs, yet was lower in vNav-MPRAGE sequences in the right inferior occipital ROI. It is noteworthy that this comparison may be in part affected by TI differences.

In reference to previous literature investigating vNav sequences, their improvement in cases with deliberate motion is evident (Andersen et al., 2019). Seminal work by Tisdall et al. reported that reductions in dice overlap ratios in FreeSurfer-acquired cortical and subcortical volumes between no-motion and motion acquisitions improved when vNav sequences were used for the latter (Tisdall et al., 2012). More recently, Ai et al. collected two MPRAGE and two vNav acquisitions within the same MRI session (Ai et al., 2019). Using ICCs, the authors found that reliability in median CT across FreeSurfer's 62 cortical regions, as assessed by MindBoggle, was highest within vNav sequences and lowest in MPRAGE sequences. Also, ICCs were not as sensitive to motion in vNav sequences, a phenomenon that was identified to be particularly relevant for CT, which typically rendered lower ICCs than area and volume. Finally, the authors demonstrated that the reproducibility (termed similarity in the current study) between MPRAGE and vNav-MPRAGE sequences was good (ICCs on average > 0.6) and exceeded the reproducibility between MPRAGE and MPRAGE pairs. Thus, the authors concluded that morphometric results obtained from MPRAGE sequences should be replicated using vNav-MPRAGE sequences.

While directly comparing CIVET or MAGeT-Brain and FreeSurfer is beyond the scope of the current work, previous studies have identified a lack of consensus between CIVET and FreeSurfer (Bhagwat et al., 2021), and have reported that CIVET is more sensitive than FreeSurfer to an Alzheimer's disease atrophic pattern at the mild cognitive impairment stage (Redolfi et al., 2015). In addition, according to previous work, MAGeT-Brain is reported to have greater accuracy than FreeSurfer (Makowski et al., 2018; Pipitone et al., 2014).

The  $^1\text{H}$ -MRS component of the current study demonstrated "excellent" or "good" reliability across the neurometabolites most commonly investigated within the literature, including Glu, Glx, Ins, Cho, NAA, and tNAA. Importantly, such findings were largely consistent across each segment of the Prisma Upgrade protocol (e.g. Trio-Trio, Trio-Prisma, Prisma-Prisma). That being said, amongst these neurometabolites, it deserves mention that the reliability of Cho levels increased in the Prisma-Prisma segment and the reliability of Ins levels decreased in the Trio-Prisma segment. Furthermore, the other  $^1\text{H}$ -MRS indices investigated in the current study included GABA, Gln, GSH, Lac, and NAAG levels (neurometabolites often excluded from 3T investigations not employing editing sequences (Wijtenburg et al., 2015)) as well as FWHM and SNR. While the failure to achieve statistical significance within several of the analyses concerning these indices renders their interpretation to be somewhat speculative, findings from the current work appear to suggest that the reliabilities of GABA, NAAG, and FWHM were highest in the Prisma-Prisma component, the reliabilities of Gln levels and SNR decreased in the Trio-Prisma component, and the reliabilities of GSH and Lac levels were highest in the Trio-Trio component. It is noteworthy that, since the RF pulse waveforms and pulse durations for the SPECIAL MRS sequence were identical on both Tim Trio and Prisma scanner implementations, the chemical shift displacements were also identical. Nonetheless, the  $^1\text{H}$ -MRS findings from the current study necessitate consideration of the protocol parameter changes between Tim Trio and Prisma acquisitions.

The present work should be considered in light of several limitations. First, the present work aimed to standardize pre- and post-scans and thus does not comprehensively assess the full benefits of the Prisma upgrade. Second, the potential influence of differences between imaging parameters between pre- and post-scans as well as between MPRAGE and vNav-MPRAGE sequences cannot be discounted. Third, the advantages of vNavs may be more apparent in instances with marked move-

ment, such as in populations that are more prone to movement or cases of forced movement. Indeed, prospective motion correction using vNavs may have a negative impact on image quality in cases where participants are relatively still. In the current work, motion was not directly assessed. Overall, the current work examines a participant sample that may be unrepresentative of the clinical population in whom prospective motion correction is beneficial; rather, it is likely representative of the control sample in large-scale case-control studies. Fourth, the results of the present study do not necessarily generalize to other sites performing similar scanner upgrades. To this effect, previous work has reported upon the between-scanner variability that exists within a scanner model (Fox et al., 2012; Takao et al., 2013), which contributes towards limiting the generalizability of our conclusions; additional work involving other scanners is undoubtedly necessitated. We see this as a first step towards obtaining this type of generalizability. As neuroimaging studies continue, merging data collected with novel techniques with upgraded hardware will become a significant challenge. Fifth, the SNR analysis employed in the current work assumes a stationary noise distribution. While this assumption requires further investigation, SNR, as it was calculated in the present study, may be influenced by non-SNR features of the images that change between software versions. Further, phantom data was not used to confirm SNR gain. Sixth, the present study may be limited by the adopted statistical approach. Future work without missing data should consider alternative methodology, such as the permutation approach adopted by Maclaren and colleagues (Maclaren et al., 2014), to better partition sources of variance. Finally, given the test-retest reliability nature of this investigation, exclusion of outputs from analysis on the basis of quality was not performed in an effort not to bias comparisons.

## 5. Conclusions

Taken together, as part of its Prisma Upgrade protocol, the current study investigated the impact that the Prisma upgrade may have on common MRI outputs (e.g. CT, SV, <sup>1</sup>H-MRS indices). High reliability was identified in most of the MRI outputs that were investigated across the Prisma upgrade. However, suboptimal reliability and differences were observed for CT or SV estimates within certain ROIs across the Prisma upgrade. Further, as assessed by CIVET and FreeSurfer, CT increases in frontal, temporal, and cingulate ROIs, along with CT decreases in parietal ROIs, may ensue across a Prisma upgrade, which may be linked to increases in SNR, changes in protocol parameters, or other factors related to the scanner upgrade. Additionally, as assessed by MAGeT-Brain and FreeSurfer, volume decreases in the amygdala, globus pallidus, hippocampus, and striatum may result from a Prisma upgrade, along with volume increases in the thalamus; however, as above, changes in protocol parameters and other upgrade-related factors cannot be discounted. These findings should be particularly considered in the design of longitudinal study designs integrating measures from before and after the upgrade, and are also highly relevant to large-scale initiatives incorporating data from multiple sites. It would be the authors' suggestion for scanner upgrade to be considered as a covariate (i.e. random effect) in statistical analyses.

In addition, as part of its vNav Protocol, the current study reported upon ROIs within which CT and SV estimates had suboptimal reliability or differed between MPRAGE and vNav-MPRAGE sequences. As assessed by CIVET, several ROIs within which MPRAGE estimates exceeded vNav-MPRAGE estimates were identified, predominantly involving temporal areas. In contrast, as assessed by FreeSurfer, higher CT estimates in vNav-MPRAGE acquisitions were identified, including mostly parietal and frontal ROIs. It is worth mentioning that the key advantages of vNav sequences may be more apparent in instances with marked head movement, such as populations that are more prone to head movement (e.g. children, certain patient populations) or cases of forced head movement. Thus, such investigations necessitate further examination.

## Credit author statement

Eric Plitman: Conceptualization, Methodology, Formal analysis, Project administration, Writing – original draft. Aurelie Bussy: Methodology, Formal analysis, Project administration, Writing – review & editing. Vanessa Valiquette: Project administration, Writing – review & editing. Alyssa Salaciak: Project administration, Writing – review & editing. Raihaan Patel: Formal analysis, Writing – review & editing. Lani Cupo: Formal analysis, Writing – review & editing. Marie-Lise Béland: Project administration, Writing – review & editing. Stephanie Tullo: Formal analysis, Writing – review & editing. Christine Tardif: Methodology, Writing – review & editing. M. Natasha Rajah: Funding acquisition, Methodology, Writing – review & editing. Jamie Near: Methodology, Formal analysis, Writing – review & editing. Gabriel A. Devenyi: Methodology, Formal analysis, Writing – review & editing. M. Malar Chakravarty: Conceptualization, Funding acquisition, Methodology, Formal analysis, Project administration, Resources, Supervision, Writing – original draft.

## Declaration of Competing Interest

Authors report no conflicts of interest.

## Acknowledgements

Support from the CIC and Douglas Research Centre and its staff who supported this work as well as the FRQS PSRv2 Grant and HBHL Platform Grant that supported the Prisma-Fit upgrade at the Douglas CIC. Data processing was partially performed using the gpc resources of the SciNet HPC Consortium.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2021.118172](https://doi.org/10.1016/j.neuroimage.2021.118172).

## References

- Ai, L., Craddock, R.C., Tottenham, N., Dyke, J.P., Colcombe, S., Milham, M., Franco, A.R., 2019. Is it time to switch your T1W sequence? Assessing the impact of prospective motion correction on the reliability and quality of structural imaging. *bioRxiv* doi:10.1101/666289.
- Alexander-Bloch, A., Clasen, L., Stockman, M., Ronan, L., Lalonde, F., Giedd, J., Raznahan, A., 2016. Subtle in-scanner motion biases automated measurement of brain anatomy from in vivo MRI. *Hum. Brain Mapp.* 37, 2385–2397.
- Andersen, M., Björkman-Burtscher, I.M., Marsman, A., Petersen, E.T., Boer, V.O., 2019. Improvement in diagnostic quality of structural and angiographic MRI of the brain using motion correction with interleaved, volumetric navigators. *PLoS One* 14, e0217145.
- Bedford, S.A., Park, M.T.M., Devenyi, G.A., Tullo, S., Germann, J., Patel, R., Anagnostou, E., Baron-Cohen, S., Bullmore, E.T., Chura, L.R., Craig, M.C., Ecker, C., Floris, D.L., Holt, R.J., Lenroot, R., Lerch, J.P., Lombardo, M.V., Murphy, D.G.M., Raznahan, A., Ruigrok, A.N.V., Smith, E., Spencer, M.D., Suckling, J., Taylor, M.J., Thurm, A., Consortium, MRC AIMS, Lai, M.-C., Chakravarty, M.M., 2020. Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. *Mol. Psychiatry* 25, 614–628.
- Bhagwat, N., Barry, A., Dickie, E.W., Brown, S.T., Devenyi, G.A., Hatano, K., DuPre, E., Dagher, A., Chakravarty, M., Greenwood, C.M.T., Misis, B., Kennedy, D.N., Poline, J.-B., 2021. Understanding the impact of preprocessing pipelines on neuroimaging cortical surface analyses, 10. GigaScience doi:10.1101/2020.05.22.100180.
- Bussy, A., Plitman, E., Patel, R., Tullo, S., Salaciak, A., Bedford, S.A., Farzin, S., Béland, M.-L., Valiquette, V., Kazanian, C., Tardif, C., Devenyi, G.A., Chakravarty, M., Alzheimer's Disease Neuroimaging Initiative, 2020. Hippocampal subfield volumes across the healthy lifespan and the effects of MR sequence on estimates. *Neuroimage*.
- Cannon, T.D., Chung, Y., He, G., Sun, D., Jacobson, A., van Erp, T.G.M., McEwen, S., Addington, J., Bearden, C.E., Cadenhead, K., Cornblatt, B., Mathalon, D.H., McGlashan, T., Perkins, D., Jeffries, C., Seidman, L.J., Tsuang, M., Walker, E., Woods, S.W., Heinsen, R. North American Prodrome Longitudinal Study Consortium, 2015. Progressive reduction in cortical thickness as psychosis develops: a multisite longitudinal neuroimaging study of youth at elevated clinical risk. *Biol. Psychiatry* 77, 147–157.
- Casey, B.J., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., Soules, M.E., Teslovich, T., Dellarco, D.V., Garavan, H., Orr, C.A., Wager, T.D., Banich, M.T., Speer, N.K., Sutherland, M.T., Riedel, M.C., Dick, A.S., Bjork, J.M.,



- Thomas, K.M., Chaarani, B., Mejia, M.H., Hagler Jr, D.J., Daniela Cornejo, M., Sicut, C.S., Harms, M.P., Dosenbach, N.U.F., Rosenberg, M., Earl, E., Bartsch, H., Watts, R., Polimeni, J.R., Kuperman, J.M., Fair, D.A., Dale, A.M. ABCD Imaging Acquisition Workgroup, 2018. The Adolescent Brain Cognitive Development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* 32, 43–54.
- Chakravarty, M.M., Steadman, P., van Eede, M.C., Calcott, R.D., Gu, V., Shaw, P., Raznahan, A., Collins, D.L., Lerch, J.P., 2013. Performing label-fusion-based segmentation using multiple automatically generated templates. *Hum. Brain Mapp.* 34, 2635–2654.
- Chincarini, A., Sensi, F., Rei, L., Gemme, G., Squarcia, S., Longo, R., Brun, F., Tangaro, S., Bellotti, R., Amoroso, N., Bocchetta, M., Redolfi, A., Bosco, P., Boccardi, M., Frisoni, G.B., Nobili, F. Alzheimer's Disease Neuroimaging Initiative, 2016. Integrating longitudinal information in hippocampal volume measurements for the early detection of Alzheimer's disease. *Neuroimage* 125, 834–847.
- Cicchetti, D.V., 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* 6, 284–290.
- Coupé, P., Manjón, J.V., Gedamu, E., Arnold, D., Robles, M., Collins, D.L., 2009. An object-based method for Rician noise estimation in MR images. *Med. Image Comput. Comput. Assist. Interv.* 12, 601–608.
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9, 179–194.
- de Graaf, R.A., 2019. *In Vivo NMR Spectroscopy: Principles and Techniques*. John Wiley & Sons.
- Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355.
- Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9, 195–207.
- Fjell, A.M., Grydeland, H., Krogstad, S.K., Amlien, I., Rohani, D.A., Ferschmann, L., Storsve, A.B., Tamnes, C.K., Sala-Llonch, R., Due-Tønnessen, P., Bjørnerud, A., Solsnes, A.E., Håberg, A.K., Skranes, J., Bartsch, H., Chen, C.-H., Thompson, W.K., Panizzon, M.S., Kremen, W.S., Dale, A.M., Walhovd, K.B., 2015. Development and aging of cortical thickness correspond to genetic organization patterns. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15462–15467.
- Fox, R.J., Sakaie, K., Lee, J.-C., Debbins, J.P., Liu, Y., Arnold, D.L., Melhem, E.R., Smith, C.H., Philips, M.D., Lowe, M., Fisher, E., 2012. A validation study of multicenter diffusion tensor imaging: reliability of fractional anisotropy and diffusivity values. *AJNR Am. J. Neuroradiol.* 33, 695–700.
- Green, P., MacLeod, C.J., 2016. SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods Ecol. Evol.* doi:10.1111/2041.12504.
- Gruetter, R., Tkáč, I., 2000. Field mapping without reference scan using asymmetric echo-planar techniques. *Magn. Reson. Med.* 43, 319–323.
- Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., Maguire, P., Rosas, D., Makris, N., Dale, A., Dickerson, B., Fischl, B., 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. *Neuroimage* 32, 180–194.
- L Jack Jr, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L.G., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27, 685–691.
- Jack, C.R., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., Shaw, L.M., Vemuri, P., Wiste, H.J., Weigand, S.D., Lesnick, T.G., Pankratz, V.S., Donohue, M.C., Trojanowski, J.Q., 2013. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* 12, 207–216.
- Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., Macfall, J., Fischl, B., Dale, A., 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage* 30, 436–443.
- Jovicich, J., Czanner, S., Han, X., Salat, D., van der Kouwe, A., Quinn, B., Pacheco, J., Albert, M., Killiany, R., Blacker, D., Maguire, P., Rosas, D., Makris, N., Gollub, R., Dale, A., Dickerson, B.C., Fischl, B., 2009. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *Neuroimage* 46, 177–192.
- Klein, A., Tourville, J., 2012. 101 labeled brain images and a consistent human cortical labeling protocol. *Front. Neurosci.* 6, 1–12.
- Lee, H., Nakamura, K., Narayanan, S., Brown, R.A., Arnold, D.L. Alzheimer's Disease Neuroimaging Initiative, 2019. Estimating and accounting for the effect of MRI scanner changes on longitudinal whole-brain volume change measurements. *Neuroimage* 184, 555–565.
- Lerch, J.P., Evans, A.C., 2005. Cortical thickness analysis examined through power analysis and a population simulation. *Neuroimage* 24, 163–173.
- Maclaren, J., Han, Z., Vos, S.B., Fischbein, N., Bammer, R., 2014. Reliability of brain volume measurements: a test-retest dataset. *Sci Data* 1, 140037.
- Makowski, C., Béland, S., Kostopoulos, P., Bhagwat, N., Devenyi, G.A., Malla, A.K., Joobar, R., Lepage, M., Chakravarty, M.M., 2018. Evaluating accuracy of striatal, pallidum, and thalamic segmentation methods: Comparing automated approaches to manual delineation. *Neuroimage* 170, 182–198.
- Mekle, R., Mlynárik, V., Gamberato, G., Hergt, M., Krueger, G., Gruetter, R., 2009. MR spectroscopy of the human brain with enhanced signal intensity at ultrashort echo times on a clinical platform at 3T and 7T. *Magn. Reson. Med.* 61, 1279–1285.
- Pardoe, H.R., Kucharsky Hiess, R., Kuzniecky, R., 2016. Motion and morphometry in clinical and nonclinical populations. *Neuroimage* 135, 177–185.
- Pipitone, J., Park, M.T.M., Winterburn, J., Lett, T.A., Lerch, J.P., Pruessner, J.C., Lepage, M., Voineskos, A.N., Chakravarty, M.M. Alzheimer's Disease Neuroimaging Initiative, 2014. Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. *Neuroimage* 101, 494–512.
- Potvin, O., Khademi, A., Chouinard, I., Farokhian, F., Dieumegarde, L., Leppert, L., Hoge, R., Rajah, M.N., Bellec, P., Duchesne, S.CIMA-Q, de Cubicaray, G., 2019. Measurement variability following MRI system upgrade. *Front. Neurol.* 10, 726.
- Provencher, S.W., 2001. Automatic quantitation of localized in vivo <sup>1</sup>H spectra with LCModel. *Neuroimaging Biomed* 14, 260–264.
- Raznahan, A., Shaw, P.W., Lerch, J.P., Clasen, L.S., Greenstein, D., Berman, R., Pipitone, J., Chakravarty, M.M., Giedd, J.N., 2014. Longitudinal four-dimensional mapping of subcortical anatomy in human development. *Proc. Natl. Acad. Sci. U. S. A.* 111, 1592–1597.
- Reardon, P.K., Seidlitz, J., Vandekar, S., Liu, S., Patel, R., Park, M.T.M., Alexander-Bloch, A., Clasen, L.S., Blumenthal, J.D., Lalonde, F.M., Giedd, J.N., Gur, R., Gur, R., Lerch, J.P., Chakravarty, M.M., Satterthwaite, T., Shinohara, R.T., Raznahan, A., 2018. Normative brain size variation and brain shape diversity in humans. *Science* 360, 1222–1227.
- Redolfi, A., Manset, D., Barkhof, F., Wahlund, L.-O., Glatard, T., Mangin, J.-F., Frisoni, G.B. neuGRID Consortium, for the Alzheimer's Disease Neuroimaging Initiative, 2015. Head-to-head comparison of two popular cortical thickness extraction algorithms: a cross-sectional and longitudinal study. *PLoS One* 10, e0117692.
- Reuter, M., Tisdall, M.D., Qureshi, A., Buckner, R.L., van der Kouwe, A.J.W., Fischl, B., 2015. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *Neuroimage* 107, 107–115.
- Schmaal, L., Veltman, D.J., van Erp, T.G.M., Sämann, P.G., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W.J., Vernooij, M.W., Ikram, M.A., Wittfeld, K., Grabe, H.J., Block, A., Hegenscheid, K., Völzke, H., Hoehn, D., Czisch, M., Lagopoulos, J., Hatton, S.N., Hickie, I.B., Goya-Maldonado, R., Krämer, B., Gruber, O., Couvy-Duchesne, B., Rentería, M.E., Strike, L.T., Mills, N.T., de Zubicaray, G.I., McMahon, K.L., Medland, S.E., Martin, N.G., Gillespie, N.A., Wright, M.J., Hall, G.B., MacQueen, G.M., Frey, E.M., Carballo, A., van Velzen, L.S., van Tol, M.J., van der Wee, N.J., Veer, I.M., Walter, H., Schnell, K., Schramm, E., Normann, C., Schoepf, D., Konrad, C., Zurowski, B., Nickson, T., McIntosh, A.M., Pappmeyer, M., Whalley, H.C., Sussman, J.E., Godlewska, B.R., Cowen, P.J., Fischer, F.H., Rose, M., Penninx, B.W.J.H., Thompson, P.M., Hibar, D.P., 2016. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. *Mol. Psychiatry* 21, 806–812.
- Shuter, B., Yeh, I.B., Graham, S., Au, C., Wang, S.-C., 2008. Reproducibility of brain tissue volumes in longitudinal studies: effects of changes in signal-to-noise ratio and scanner software. *Neuroimage* 41, 371–379.
- Siemens, 2020. MAGNETOM Trio Upgrade. URL <https://www.siemens-healthineers.com/magnetic-resonance-imaging/options-and-upgrades/upgrades/magnetom-trio-upgrade/features>
- Simpson, R., Devenyi, G.A., Jezzard, P., Hennessy, T.J., Near, J., 2017. Advanced processing and simulation of MRS data using the FID appliance (FID-A)-An open source, MATLAB-based toolkit. *Magn. Reson. Med.* 77, 23–33.
- Takao, H., Hayashi, N., Ohtomo, K., 2013. Effects of the use of multiple scanners and of scanner upgrade in longitudinal voxel-based morphometry studies. *J. Magn. Reson. Imaging* 38, 1283–1291.
- Tisdall, M.D., Dylan Tisdall, M., Hess, A.T., Reuter, M., Meintjes, E.M., Fischl, B., van der Kouwe, A.J.W., 2012. Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magn. Reson. Med.* 68, 389–399.
- Tisdall, M.D., Reuter, M., Qureshi, A., Buckner, R.L., Fischl, B., van der Kouwe, A.J.W., 2016. Prospective motion correction with volumetric navigators (vNavs) reduces the bias and variance in brain morphometry induced by subject motion. *Neuroimage* 127, 11–22.
- Tkáč, I., Starcuk, Z., Y. Choi, I., Gruetter, R., 1999. In vivo <sup>1</sup>H NMR spectroscopy of rat brain at 1 ms echo time. *Magn. Reson. Med.* 41, 649–656.
- Tullo, S., Patel, R., Devenyi, G.A., Salaciak, A., Bedford, S.A., Farzin, S., Wlodarski, N., Tardif, C.L., Breitner, J.C.S., Mallar Chakravarty, M. the PREVENT-AD Research Group, 2019. MR-based age-related effects on the striatum, globus pallidus, and thalamus in healthy individuals across the adult lifespan. *Hum. Brain Mapp.* 40, 5269–5288.
- Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* 29, 1310–1320.
- van Erp, T.G.M., Hibar, D.P., Rasmussen, J.M., Glahn, D.C., Pearlson, G.D., Andriessen, O.A., Agartz, I., Westlye, L.T., Haukvik, U.K., Dale, A.M., Melle, I., Hartberg, C.B., Gruber, O., Kraemer, B., Zilles, D., Donohoe, G., Kelly, S., McDonald, C., Morris, D.W., Cannon, D.M., Corvin, A., Machielsen, M.W.J., Koenders, L., de Haan, L., Veltman, D.J., Satterthwaite, T.D., Wolf, D.H., Gur, R.C., Gur, R.E., Potkin, S.G., Mathalon, D.H., Mueller, B.A., Preda, A., Macciardi, F., Ehrlich, S., Walton, E., Hass, J., Calhoun, V.D., Bockholt, H.J., Sponheim, S.R., Shoemaker, J.M., van Haren, N.E.M., Pol, H.E.H., Ophoff, R.A., Kahn, R.S., Roiz-Santiañez, R., Crespo-Facorro, B., Wang, L., Alpert, K.I., Jönsson, E.G., Dimitrova, R., Bois, C., Whalley, H.C., McIntosh, A.M., Lawrie, S.M., Hashimoto, R., Thompson, P.M., Turner, J.A., 2016. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Mol. Psychiatry* 21, 585.
- van Erp, T.G.M., Walton, E., Hibar, D.P., Schmaal, L., Jiang, W., Glahn, D.C., Pearlson, G.D., Yao, N., Fukunaga, M., Hashimoto, R., Okada, N., Yamamori, H., Bustillo, J.R., Clark, V.P., Agartz, I., Mueller, B.A., Cahn, W., de Zwarte, S.M.C., Hulshoff Pol, H.E., Kahn, R.S., Ophoff, R.A., van Haren, N.E.M., Andriessen, O.A., Dale, A.M., Doan, N.T., Gurholt, T.P., Hartberg, C.B., Haukvik, U.K., Jørgensen, K.N., Lagerberg, T.V., Melle, I., Westlye, L.T., Gruber, O., Kraemer, B., Richter, A.,

- Zilles, D., Calhoun, V.D., Crespo-Facorro, B., Roiz-Santiañez, R., Tordesillas-Gutiérrez, D., Loughland, C., Carr, V.J., Catts, S., Croy, V.L., Fullerton, J.M., Green, M.J., Henskens, F.A., Jablensky, A., Lenroot, R.K., Mowry, B.J., Michie, P.T., Pantelis, C., Quidé, Y., Schall, U., Scott, R.J., Cairns, M.J., Seal, M., Tooney, P.A., Rasser, P.E., Cooper, G., Shannon Weickert, C., Weickert, T.W., Morris, D.W., Hong, E., Kochunov, P., Beard, L.M., Gur, R.E., Gur, R.C., Satterthwaite, T.D., Wolf, D.H., Belger, A., Brown, G.G., Ford, J.M., Macciardi, F., Mathalon, D.H., O'Leary, D.S., Potkin, S.G., Preda, A., Voyvodic, J., Lim, K.O., McEwen, S., Yang, F., Tan, Y., Tan, S., Wang, Z., Fan, F., Chen, J., Xiang, H., Tang, S., Guo, H., Wan, P., Wei, D., Bockholt, H.J., Ehrlich, S., Wolthuisen, R.P.F., King, M.D., Shoemaker, J.M., Sponheim, S.R., De Haan, L., Koenders, L., Machielsen, M.W., van Amelsvoort, T., Veltman, D.J., Assogna, F., Banaj, N., de Rossi, P., Iorio, M., Piras, F., Spalletta, G., McKenna, P.J., Pomarol-Clotet, E., Salvador, R., Corvin, A., Donohoe, G., Kelly, S., Whelan, C.D., Dickie, E.W., Rotenberg, D., Voineskos, A.N., Ciufolini, S., Radua, J., Dazzan, P., Murray, R., Reis Marques, T., Simmons, A., Borgwardt, S., Egloff, L., Harrisberger, F., Riecher-Rössler, A., Smieskova, R., Alpert, K.L., Wang, L., Jönsson, E.G., Koops, S., Sommer, I.E.C., Bertolino, A., Bonvino, A., Di Giorgio, A., Neilson, E., Mayer, A.R., Stephen, J.M., Kwon, J.S., Yun, J.-Y., Cannon, D.M., McDonald, C., Lebedeva, I., Tomyshev, A.S., Akhadov, T., Kaleda, V., Fatouros-Bergman, H., Flyckt, L., Karolinska Schizophrenia Project, Busatto, G.F., Rosa, P.G.P., Serpa, M.H., Zanetti, M.V., Hoschl, C., Skoch, A., Spaniel, F., Tomecek, D., Hagenaars, S.P., McIntosh, A.M., Whalley, H.C., Lawrie, S.M., Knöchel, C., Oertel-Knöchel, V., Stäblein, M., Howells, F.M., Stein, D.J., Temmingh, H.S., Uhlmann, A., Lopez-Jaramillo, C., Dima, D., McMahon, A., Faskowitz, J.L., Gutman, B.A., Jahanshad, N., Thompson, P.M., Turner, J.A., 2018. Cortical Brain Abnormalities in 4474 Individuals With Schizophrenia and 5098 Control Subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. *Biol. Psychiatry* 84, 644–654.
- van Haren, N.E.M., Schnack, H.G., Cahn, W., van den Heuvel, M.P., Lepage, C., Collins, L., Evans, A.C., Hulshoff Pol, H.E., Kahn, R.S., 2011. Changes in cortical thickness during the course of illness in schizophrenia. *Arch. Gen. Psychiatry* 68, 871–880.
- van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G.F., Calderoni, S., Daly, E., Deruelle, C., Di Martino, A., Dinstein, I., Duran, F.L.S., Durston, S., Ecker, C., Fair, D., Fedor, J., Fitzgerald, J., Freitag, C.M., Gallagher, L., Gori, I., Haar, S., Hoekstra, L., Jahanshad, N., Jalbrzikowski, M., Janssen, J., Lerch, J., Luna, B., Martinho, M.M., McGrath, J., Muratori, F., Murphy, C.M., Murphy, D.G.M., O'Hearn, K., Oranje, B., Paredes, M., Retico, A., Rosa, P., Rubia, K., Shook, D., Taylor, M., Thompson, P.M., Tosetti, M., Wallace, G.L., Zhou, F., Buitelaar, J.K., 2018. Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. *Am. J. Psychiatry* 175, 359–369.
- Voineskos, A.N., Winterburn, J.L., Felsky, D., Pipitone, J., Rajji, T.K., Mulsant, B.H., Chakravarty, M.M., 2015. Hippocampal (subfield) volume and shape in relation to cognitive performance across the adult lifespan. *Hum. Brain Mapp.* 36, 3020–3037.
- Wijtenburg, S.A., Yang, S., Fischer, B.A., Rowland, L.M., 2015. In vivo assessment of neurotransmitters and modulators with magnetic resonance spectroscopy: application to schizophrenia. *Neurosci. Biobehav. Rev.* 51, 276–295.