Multi-Site Structural Magnetic Resonance Imaging of Myelin

By:

Laagishan Yoganathan BSc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University

Master of Science (2019) McMaster University
Department of Psychology Neuroscience & Behaviour Hamilton, Ontario

Title: Multi-Site Structural MRI for Cortical Myelin Mapping

Author: Laagishan Yoganathan
Supervisor: Dr. Nicholas Bock

Number of Pages: vi, 32

Abstract

Multi-site MRI studies collect large amounts of data in a short time frame. Large sample sizes are desirable to address power and replicability issues that have been problematic for scientists in the past. Although multi-site MRI solves the sample size problem, it brings with it a new set of challenges. Scanning the same person at different sites might result in differences in MRI derived measurements. In this thesis we compared three approaches to facilitate the analysis of multi-site MRI data: quantitative R_1 mapping, adding site as a covariate in a linear model, and using the ComBat method. We also investigated the relationship between two common MRI measurements: signal and volume. We collected data from 64 healthy participants across 3 GE scanners and 1 Siemens scanner at 3T. We found that signal intensity was different between vendors whereas volume was not. Our R_1 method resulted in values that were different across vendor and significantly lower than those reported in the literature. B_1+ maps used to calculate R_1 were different across sites. Using a scale factor, we were able to compensate for mistakes in R_1 mapping. We also found that adding site as a covariate corrected mean differences in signal intensity across sites, but not differences in variance. The ComBat method gave best similarity between sites. However, since different people were scanned at each site, we couldn't evaluate the effectiveness of each method as variation in the data could have been due to site effects or heterogeneity in participants. White matter volume and signal intensity in the white matter were correlated in males but not in females. We found that this low correlation was caused by outliers in our female sample. The correlation between white matter volume and signal in males suggests that both metrics are measuring myelin and can be used as converging evidence to detect changes in brain myelination.

Acknowledgements

The person I want to thank the most is my thesis supervisor, Dr. Nicholas Bock. For taking me on as a student, and paving the way for my career, I will be forever grateful.

Table of Contents

1. Magnetic Resonance Imaging for Myelin

The brain is highly complex yet highly organized. Modelling this organization could improve our understanding of structure-function relationships in the brain. There is a lot of interest in understanding the information processing capabilities of the brain, and the region where most information processing takes place is the cerebral cortex. It is the 2-4mm thick outer layer of the brain which contains billions of neurons and even more synapses and axons (Sporns, Tononi, & Kötter, 2005). It is responsible for several functions including speech, movement, visual perception, attention and working memory. Many tools are used to study cortical structure and organization including magnetic resonance imaging (MRI), histology, and electrophysiology.

The organization of neuronal cell bodies in the cortex is called cytoarchitecture. In general, the cortex is organized into 6 layers, although in some regions it is difficult to see distinctions between layers. Layer I has few cells and mostly consists of dendrites and axons. Layer II and IV are mostly made up of granule cells. Layer III and V are mostly made up of pyramidal neurons. Layer VI is made up of a variety of neurons. The composition of each layer varies regionally. For example, primary motor cortex has a smaller layer IV and a larger layer V because of large pyramidal cells in layer V of the motor cortex. The 52 Broadmann areas were delineated based on this difference in cytoarchitecture.

Another aspect of cortical architecture is myeloarchitecutre, the distribution of myelinated axons in the cortex which also has a layered pattern. Layers I-III mostly consist of intracortical connections. Layer IV neurons receive input from thalamus. Layer V neurons connect down to the brainstem and spinal cord while Layer VI sends connections to the thalamus. Myeloarchitecutre also varies regionally. For example, the visual cortex has a large, heavily myelinated layer IV (input layer) which forms the Stria of Gennari.

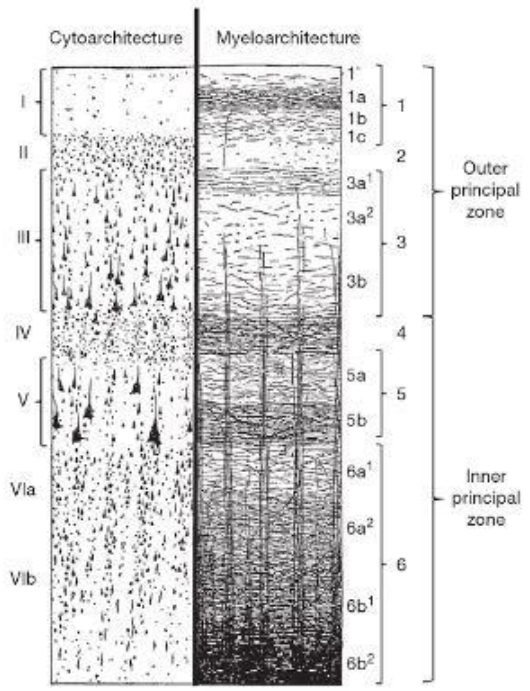


Figure 1: Taken from Nieuwenhuys, 2013. Figure 3 in their paper. This shows how both cytoarchitecture and myeloarchitecture have a layered pattern in the cortex. Permission obtained from Springer.

The fine details of cytoarchitecture and myeloarchitecutre are studied using histology stains. Although histology can be used to study the brain at micrometer resolution, with MRI it is possible to study certain properties of the brain in 3D and *in vivo*, whereas histology is usually performed on 2D slices of post-mortem tissue. Structural MRI can be used to quantify properties of the cortex such as cortical thickness, gyrification, surface area & myeloarchitecture. Several MRI methods have been developed to investigate myeloarchitecture across the entire cortex. Although studies have shown the possibility of investigating layer specific myelin fibers of the cortex using *ex-vivo* diffusion MRI (Leuze et al., 2012), more common imaging techniques used in clinical studies are not able to produce images at such fine resolution. In most structural MRI studies, myelin is summarized using a spatially averaged MRI metric (such as longitudinal relaxation rate ($R_1 = \frac{1}{T_1}$), T_1-weighted/proton density-weighted ratio, T_1-weighted/T_2-weighted ratio) across the depths of the cortex. A "myelin map" of the cortex using MRI shows that primary sensory cortical areas are more heavily myelinated than association areas like frontal and parietal cortices.

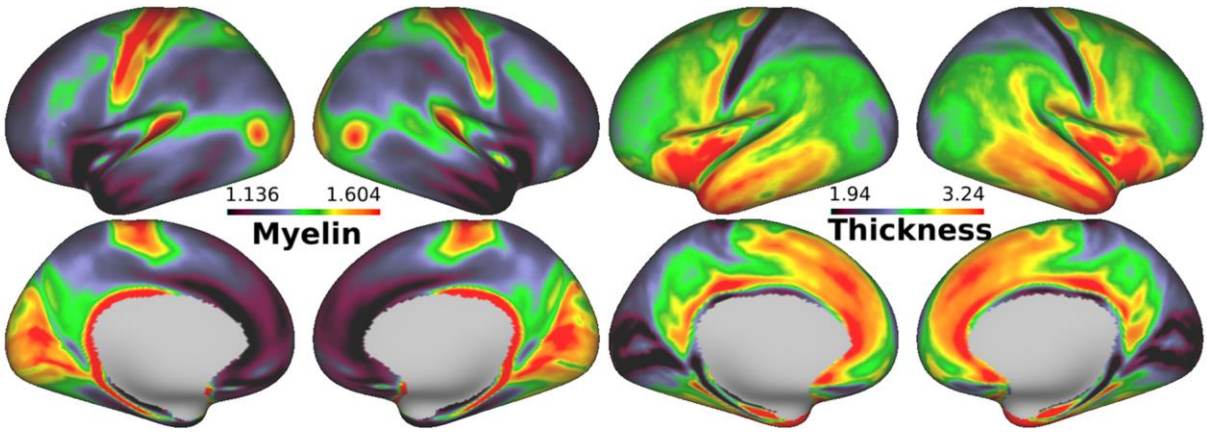


Figure 2: Taken from Fukutomi et. al, 2018. Part of figure 1 from their paper. This shows average cortical myelin and average cortical thickness in a sample of 1000 subjects from the Human Connectome Project. Myelin in the figure refers to intensity in a T_1-weighted/T_2-weighted ratio image. Open access CC-BY licence on Elsevier.

Another advantage of MRI is that it can be used to study lifespan trajectories of brain development. With MRI, it is possible to collect data from many living people allowing for large scale cross sectional studies and the ability to track brain changes in individuals longitudinally. Lifespan trajectories can be used to study healthy development, disease progression and the effect of interventions on the brain. Disease studies tend to look at mean group differences in MRI metrics, but differences in developmental trajectories can provide complimentary insight into disease pathophysiology. For example a study investigating symptom progression in individuals with Alzheimer's was able to group individuals based on trajectory of clinical decline as well as predict individual trajectories using clinical scores and MRI metrics (Bhagwat et al., 2018). Another study investigating the relationship between cortical thickness and IQ in children found that there were modest non-significant correlations between IQ and thickness which varied with age. Splitting children into different groups based on IQ, they found that differences in developmental trajectory of cortical thickness separated children with different levels of intelligence (Shaw et al., 2006).

In terms of healthy brain aging, grey matter development peaks early and declines throughout life, whereas white matter peaks later in life. Histological myelin staining and MRI derived measures of myelin both increase until middle adulthood and then decline. A large scale MRI study of adulthood trajectory of intracortical myelin showed it plateaus in the late 30s before beginning to decline in in the 50s (Grydeland et al., 2013). This trajectory seems to be affected in mental health disorders such as bipolar disorder (Sehmbi et al., 2019) and schizophrenia (Tishler et al., 2018). One of the drawbacks of aging research is the larger sample size required to detect effects. Larger sample sizes are needed for detecting differences in trends between groups in cross-sectional data and even larger samples are required when covariates are considered (Naiji et al., 2013).

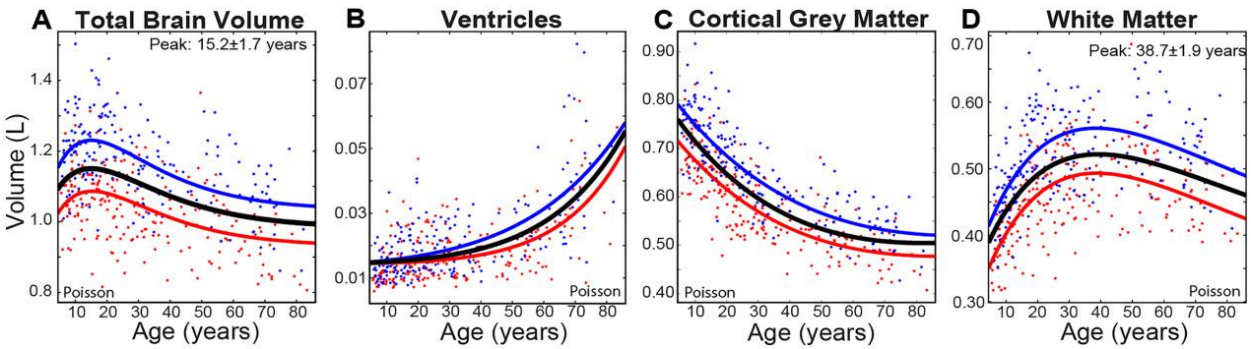


Figure 3: Taken from Narvacan et. al, 2018. Figure 2 from their paper. This shows lifespan trajectories for some metrics obtainable from MRI in a cross sectional healthy population. Blue lines represent males and red lines represent females, while black is the entire data. Permission obtained from Wiley.

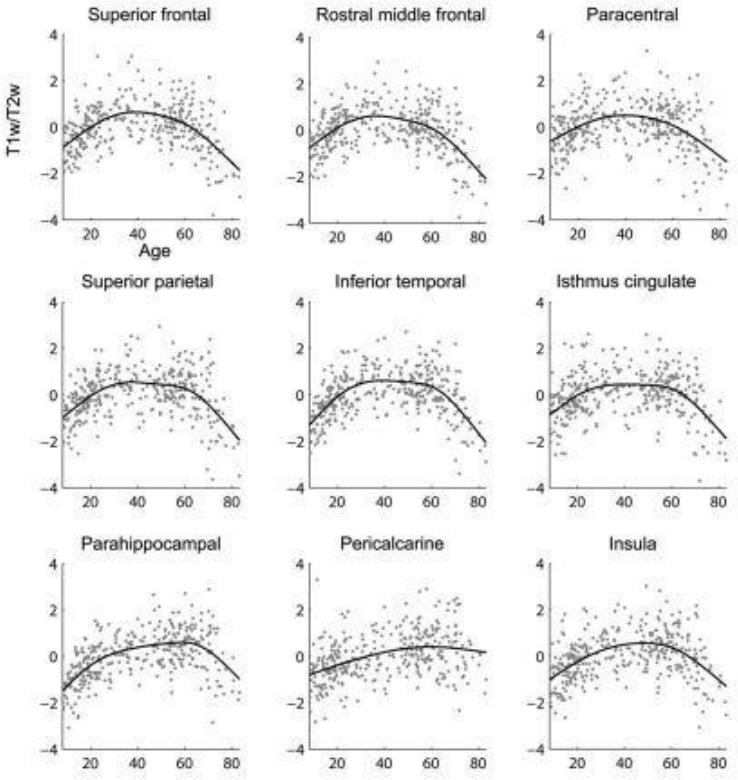


Figure 4: Taken from Grydeland et al., 2013. Figure 6 from their paper. This shows intracortical myelin follows an inverted U trajectory in healthy aging. Open access CC-BY licence from Journal of Neuroscience..

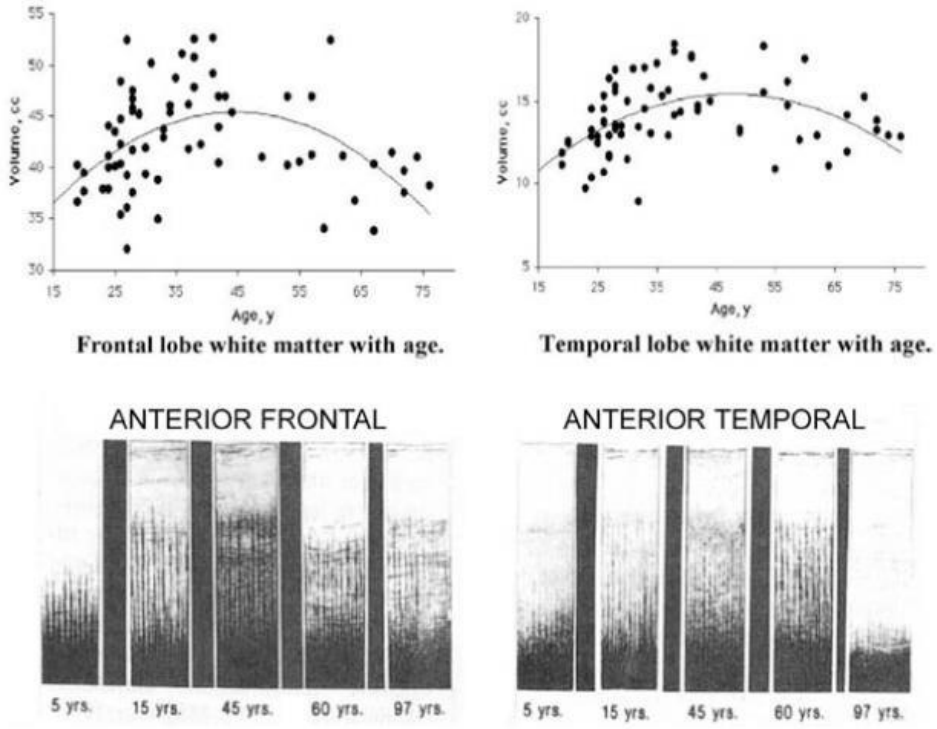


Figure 5: Taken from Bartzokis 2005. Figure 1 in their paper. This shows the relationship between myelin stains and white matter volume. NIH/HHS public access

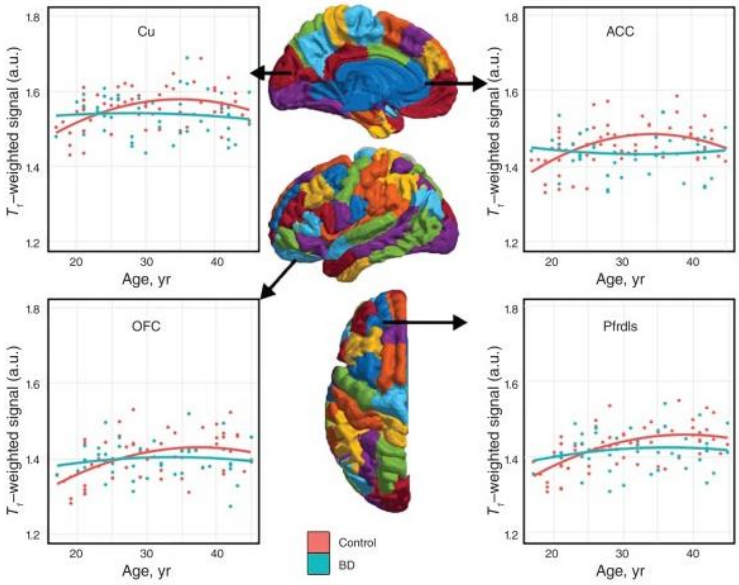


Figure 6: Taken from Sehmbi et al., 2019. This shows the age trajectory of intracortical myelin is less quadratic in BD type-I compared to healthy controls. NIH public access

2. **MRI Signal and Bias:**

Histological studies quantify myelin content by staining the lipids and proteins in myelin while MRI creates contrast using differences in tissue composition. The signal

intensity from an MRI image is dependent on properties of the tissue inside the scanner such as spin density, longitudinal relaxation time (T_1), transverse relaxation time (T_2) and chemical shift. Various types of contrast can be achieved by suppressing or enhancing the effects of these properties. By using specific imaging sequences, we can sensitize an image to a tissue property such as T_1 which would produce a T_1-weighted image (T1w). The most common T1w imaging sequence is magnetization prepared rapid gradient echo (MPRAGE), which generates strong grey matter/white matter contrast arising from differences in T_1 between grey matter and white matter. White matter has more myelin and thus more lipid molecules that shorten T_1 (faster relaxation) of nearby water. T_1 of grey matter is 1400ms and T_1 of white matter is 900 ms at 3T, although this estimate varies depending on the measurement method used (Stanisz et al., 2005; Stikov et al., 2015). The signal intensity in a T1w image is correlated with myelin observed with histology. It has been shown that the T_1 signal follows myeloarchitecture more closely then cytoarchitecture, although a combination of both fit the signal best (Eichkoff et al. 2005). Thus T_1 and the associated longitudinal relaxation rate R_1 ($1/T_1$) are widely used as a surrogate marker of myelin.

One of the caveats of MRI is that it is an indirect marker of the biological properties we are interested in studying. Both white matter volume and T1w MRI signal have been shown to be related to myelin, however, both are also affected by factors other than myelin. Although MRI signal intensity originates from the object itself, it can also be affected by radiofrequency (B_1) or static (B_0) field inhomogeneity. These changes in signal intensity are unrelated to the underlying tissue, and can bias interpretation and processing of MRI images. Many techniques have been developed to remove sources of bias from MRI images. For example, the N3 algorithm is a common preprocessing step to remove image intensity nonuniformity before MRI segmentation (Sled et al., 1998). N3 estimates a smoothly varying multiplicative model of the bias field and removes it. Images can also be corrected by collecting a separate map of the bias field. (Wang et al., 2005).

Another method that has been used to correct receive (B_1-) and transmit (B_1+) bias is by collecting a map of the bias field using a proton density weighted (PDw) FLASH image (Wang et al., 2005; Helms et al., 2008). This method has been used to generate bias free T1w images by dividing a T1w image by a PDw image (Van de Moortele et al., 2009; Marques et al., 2010). For a low flip angle FLASH image, the signal is almost proportional to B_1+ X B_1- (Helms et al., 2008). Dividing a T1w MPRAGE image by a PDw FLASH image has also been used to increase contrast in the cortex (Bock et al., 2013).

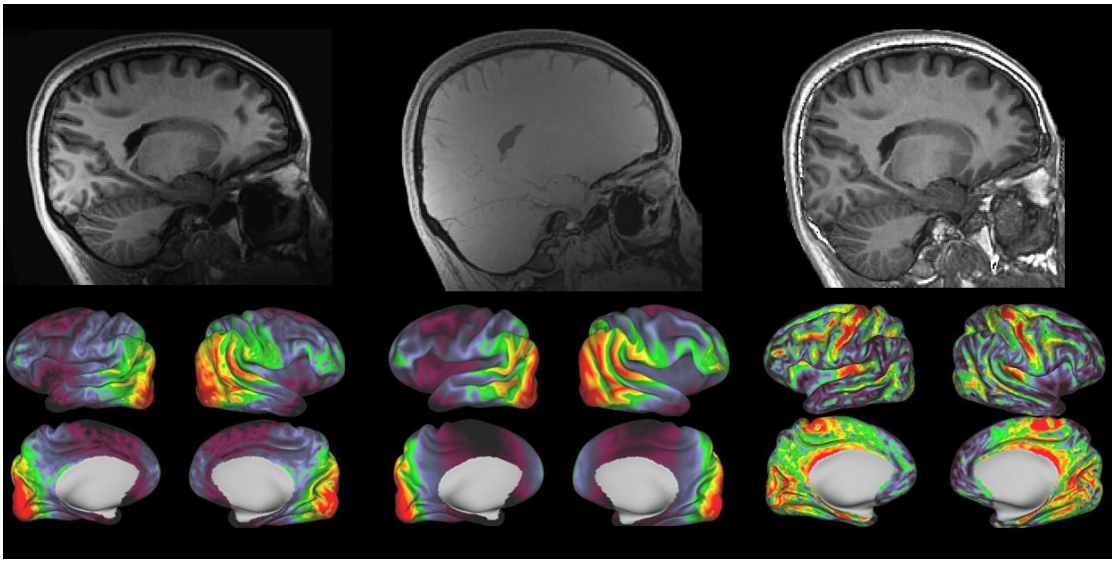


Figure 7: From left to right: T1w MPRAGE, PDw FLASH and T1w/PDw ratio images in a single subject. Signal intensity nonuniformity is visible in both the T1w and PDw images with the back of the brain being brighter due to B1+ and B1-. This is easily noticeable on the surface. The ratio image is much more uniform and on the surface we can see features of myeloarchitecture (myelinated motor and visual cortices).

3. Quantitative MRI for Multi-Site

Multi-site MRI is becoming the standard for clinical neuroimaging research in an effort to address replication and power issues which have been problematic for studies in the past. However, different sites often have different scanner vendors with a different set of pulse sequences and optimizations. Some problems of using different scanners or vendors include having different bias fields, coil geometry, or image scaling which can lead to differences in measures of cortical thickness, brain volume, white matter, & signal intensity (Chenevert et al., 2014, Haast et al., 2018). The effect of having different scanners and different protocols is a concern for the reproducibility of findings and for meta-analysis, as site and vendor specific differences can lead to systematic differences in MRI metrics.

To address this issue, rather than attempting to correct for biases and control all the variables that could affect the image at each site, we can create quantitative maps of tissue parameters that are sequence and scanner independent. Changing sequence parameters, bias fields or scanner hardware should not affect the imaged object. Quantitative MRI aims to measure properties of the underlying tissue and thus should be robust to site effects. Tissue parameters can be calculated from the signal intensity of an MRI image. The MRI signal can be described mathematically as a function of sequence parameters (such as repetition time (TR), echo time (TE), echo spacing (ES)) and tissue parameters (such as T_1, T_2*, PD). We can mathematically model signal intensity of a pulse sequence and we can calculate the value of tissue parameters given pulse sequence parameters. Using quantitative measures of tissue parameters such as R_1 might improve comparability of MRI derived metrics across sites. Recently, studies have investigated if quantitative MRI increases reliability between sites and have found

good agreement between sites, although there is still room for improvement. Lee et al., (2019) used the variable flip angle method to map T_1 and observed an intra-scanner test-retest variability of <3.8% and inter-scanner test-retest reliability that varied from 3.8% to 13.2%.

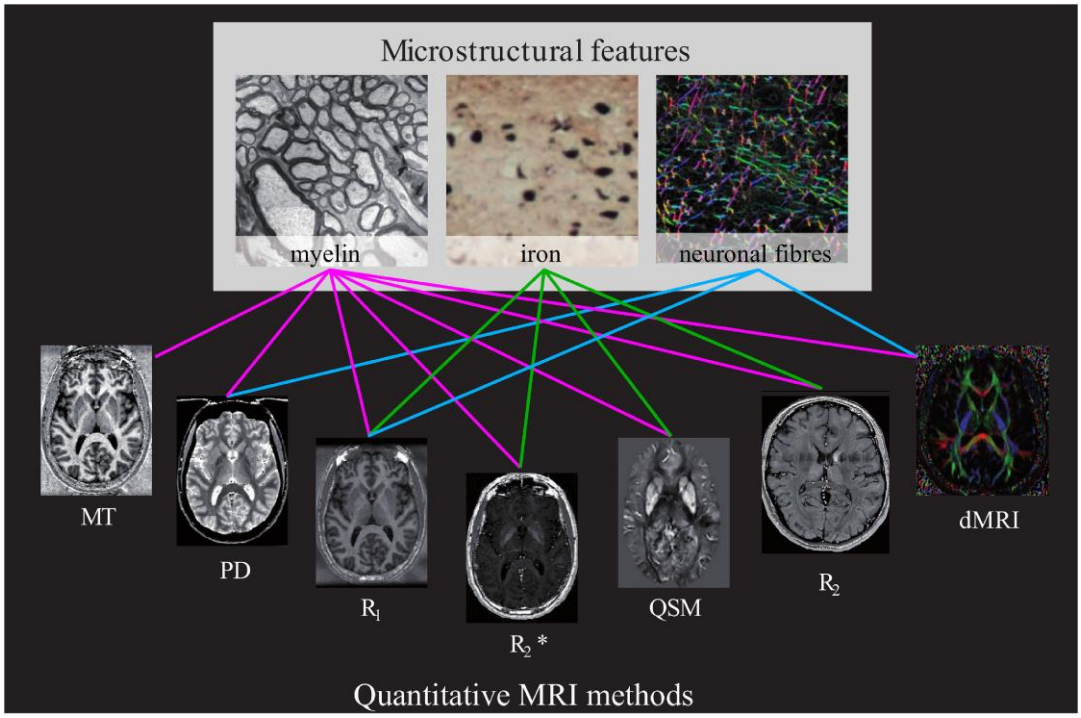


Figure 8: Taken from Edwards, Kirilina, Mohammadi & Weiskopf, (2018). Figure 3 in their paper. Different quantitative metrics can be used to investigate different aspects of cortical microstructure. Open Access CC-BY licence on Elsvier

Although other quantitative metrics exist, such as magnetization transfer which may be more specific to myelin, this thesis focuses on using R1 for myelin mapping. The main advantage of R_1 is that methods for mapping R_1 are generally faster than other methods, making it suitable for clinical studies. The disadvantages of R_1 are that estimates vary widely in the literature and R_1 is not specific to myelin. Several methods exist for quantitative R_1 mapping such as variable flip angle, MP2RAGE and inversion recovery. Tsialios et al., (2017) compared several methods and showed that R_1 values differ depending on the method used. This has led to a wide range of estimates for R_1 in grey matter and white matter. Although R_1 is a quantitative metric that is more specific than a weighted image, it is also correlated with iron. Stueber et al., (2014) collected quantitative R_1 maps and also measured iron and myelin concentration (by measuring phosphorus and sulfur) in the cortex and neighbouring white matter. They showed that 64 percent of the variation in R_1 in cortical grey matter could be attributed to myelin while 36 percent is accounted for by iron. In the white matter, 90 percent of the variation in R_1 is accounted for by myelin while 10 percent is accounted for by iron. This is because iron and myelin are co-localized in cortical grey matter (Fukunaga et al., 2010). R_1 also depends on axon diameter (Harkins et al., 2016).

To estimate R_1 in the cortex we use a technique with T1w MPRAGE and PDw FLASH images (Bock et al., 2013). Using signal equations, we can create a lookup table and solve for R_1 while also taking into account local deviations in flip angle caused by B_1+ inhomogeneity in a method similar to MP2RAGE (Marques et al., 2010).

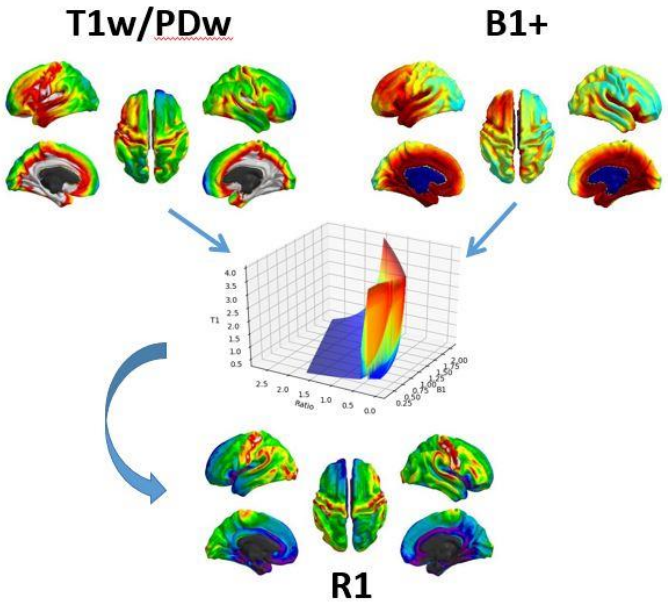
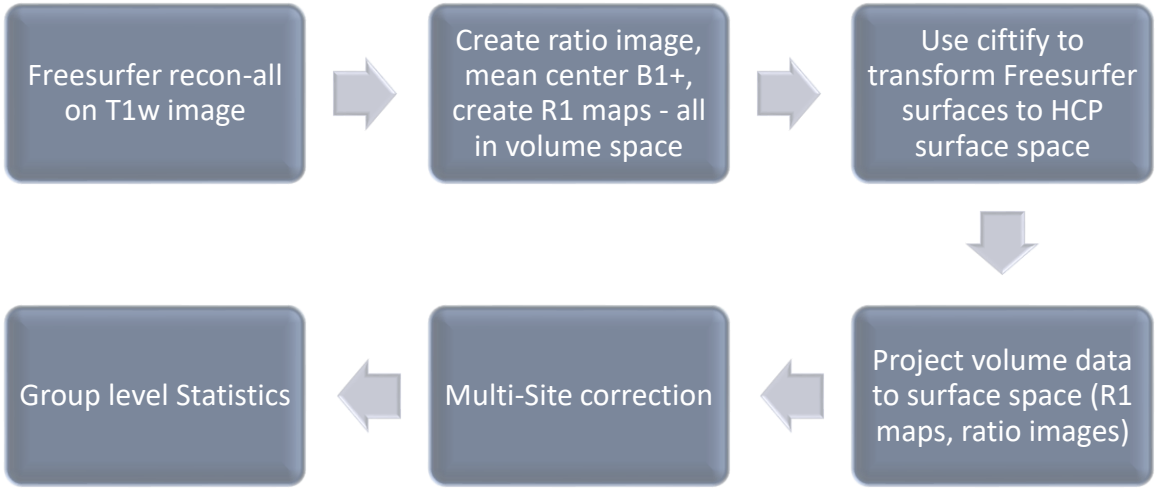


Figure 9: Our custom R1 method. A lookup table is generated from signal equations. T1w/PDw images and B1+ maps are passed and the value of R1 is interpolated from the table.

This thesis explores techniques to facilitate multi-site MRI data analysis. We show that correction for multi-site data is necessary, and correction can either be performed using an imaging protocol that is reliable across multiple sites or correction can be applied after images are acquired using statistical approaches. We collect data from four sites and calculate R_1 incorporating a scale factor for each vendor. We also test two statistical approaches: including site as a covariate in a linear model and the ComBat method (Fortin et al., 2017, Johnson, Li, & Rabinovic, 2007). We also tested the relationship between two common MRI metrics: volume and signal.

4. Methods

All data analysis was performed on Compute Canada's high performance computing cluster named Graham. Below is an outline of the MRI processing steps:



4.1 MRI Acquisition Protocols

We collected MRI data from four sites. Three had a General Electric (GE) scanners (McMaster University, Dalhousie University, and University of Calgary) and one had a Siemens scanner (Queens University). All GE sites had a 3T MR750w Discovery scanner. The Siemens site had a 3T MAGNETOM TrioTim syngo MR B19 scanner.

We collected a T1w, PDw, high contrast T1w (T1wHC) and B1+ map at each site. The T1w image is used with Freesurfer for surface generation while the T1wHC is used to create a ratio image and R_1 map. Sequence parameters for each site are listed below. TI is inversion time, TD is delay time, TE is echo time, ES is echo spacing, θ is flip angle, TR is repetition time, PE is phase encoding direction, FOV is field-of-view and TA is total acquisition time. Some parameters are calculated based on inputs into the scanner. For GE, TR is not set, but calculated via TR = TI + N*ES + TD. Similarly, for Siemens TD is not set but calculated as TD=TR-TI-N*ES.

Image	Sequence	TI(s)	TD (s)	TE (ms)	ES (ms)	θ(°)	TR (s)	PE	FOV(mm)	TA (min)
HC T1w	BRAVO	1	1.1	3.06	7.93	12	2.89	Centric	240x240x100	6:19 x 2
T1w	BRAVO	0.45	0	3.06	7.92	12	1.87	Centric	240x240x180	5:32
PDw	SPGR	-	-	3.06	7.91	4	1.42	Linear	240x240x180	4:59
B1+	2d b1map	-	-	12.3	16	15		-	64x64x35	1:29

Table 1 Sequence parameters at the 3 GE sites. TI is the time until centre of k-space is acquired. When phase encoding is centric, TI is time until the beginning of the acquisition block. The TI and TD parameters are calculated based on information input to the scanner. If opti is set to 2.1, and pos_start_ir is 1.1, then TI will be 1. For the low contrast, opti is 0.45 and pos_start_ir is 0.

Image	Sequence	TI(s)	TD(s)	TE (ms)	ES (ms)	θ(°)	TR(s)	PE	FOV(mm)	TA (min)
HC T1w	tfl	1	1.1	2.01	7.93	12	3.65	Linear	256x240x176	8:14
T1w	tfl	0.0	0.8	2.96	7.10	9	2.3	Linear	256x256x176	5:30
PDw	tfl	-	-	2.01	7.91	4	1.09	Linear	256x240x176	2:29
B1+	ep_seg_se	-	-	46	4.18	60	5	-	256x256x144	1:25
B1+	ep_seg_se	-	-	46	4.18	120	5	-	256x256x144	1:25

Table 2 Sequence parameters for the Siemens site. The TI in the acquisition protocol is to the beginning of the acquisition block. Since phase encoding is linear, k-space center is acquired in the middle of the acquisition. You need to find the time from beginning of inversion pulse to the k-space center. Prescribed TI for HC was 2100 and 900 for LC.

4.2 Subject demographics

Participants were recruited as part of an ongoing longitudinal imaging project on Type-I Bipolar Disorder. In this thesis, the focus is on the cross sectional healthy controls of this cohort (N=64).

	McMaster	Dalhousie	Calgary	Queens	Total
M(F)	14(13)	7(10)	7(1)	2(10)	30(34)
Mean Age	27.6	26.6	21.5	24.5	-

Table 3: Subject demographics for the baseline healthy controls for a longitudinal imaging project on bipolar disorder

4.3 Freesurfer for Surface Generation

We run Freesurfer v6.0 (https://surfer.nmr.mgh.harvard.edu/) recon-all pipeline on the T1w image. The T1w image has less blurring than the T1wHC image and we believe this might allow Freesurfers automatic segmentation to better identify the grey matter/white matter boundary and the grey matter/cerebrospinal fluid boundary, although this has not been formally tested.

4.4 Creating Ratio Image for Myelin Mapping

We use dcm2niix (https://www.nitrc.org/frs/?group_id=889) that is distributed with MRIcroGL to convert DICOMS to NiFTI. We used the latest version at the time of writing (release data June 23 2018).

Next, we perform image registration and create a ratio image using FSL v 5.0.11 (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). We register all images to the T1w image with linear affine registration using FLIRT and 6 degrees of freedom (DOF). We transform the Freesurfer skull-stripped brain back into native space using mri_vol2vol. We register this skull-stripped brain to the MNI152 T1 1mm brain atlas included with FSL using FLIRT and 12 DOF (atlas located in $FSLDIR/data/standard/MNI152_T1_1mm_brain.nii.gz, which corresponds to the ICBM152 Nonlinear asymmetric template 6th generation). We apply this transform to all the images – T1w, T1wHC, B1+, and PDW - which are registered to the T1w image. For GE's T1wHC, two halves were acquired separately. We register both halves to the T1w, then transform them to the MNI space using the 12 DOF FLIRT matrix that was used to register the skull stripped brain to MNI152 space.

Then we add them taking the largest value at each voxel. The end result is a whole brain high contrast image. This procedure is described in more detail elsewhere (Rowley et al., 2017). For the Siemens site, we collect a single whole brain T1wHC image and divide that by the PDw image.

4.5 B1+ Mapping and Smoothing Procedure

GE's 2db1map sequence uses Bloch-Seigert shift B_1+ mapping (Sacolick et al. 2010) and produces a B_1+ map as well as an anatomical volume. We register the B_1+ anatomical image to the T1w image using FLIRT 6 DOF. We then apply this transform to the B_1+ image. The GE B_1+ map has units of 10*flip angle. This is because the scanner cannot store decimal numbers. The prescribed flip angle was 15 degrees, so we divide the B_1+ volume values by 150 (10*15). A value of 1 would indicate the prescribed flip angle was achieved (signal intensity is a function of flip angle, so incorrect flip angle will lead to differences in signal intensity due B_1+ rather than due to differences in the underlying tissue).

On the Siemen's, we use the double angle method (Sled & Pike 2000; Wang, Qiu & Constable 2005; Wang et al. 2009). We register the 120° flip angle spin echo image to the T1w skull stripped brain from Freesurfer using FLIRT with 6 DOF. Then we register the 60° flip angle spin echo image to the 120° flip angle image. Then we use the equations for the double angle method to create a B_1+ map.

For smoothing, we multiply the B_1+ map with the brainmask generated from Freesurfer to mask out non brain areas. Then we apply a 5mm Gaussian smoothing kernel (this corresponds to mean filtering with 5mm) to the masked B_1+ map. The smoothed B_1+ map is used in R_1 calculations.

4.6 R_1 Method (Reverse Lookup Table / Dictionary method)

See here for an excellent explanation of dictionary based methods: https://qmrlab.org/2019/04/08/T1-mapping-mp2rage.html. The author of the blog also provides intuitive explanations to other quantitative MRI methods such as inversion recovery and variable flip angle.

For our method, we model the signal intensity of the ratio image (T1w/PDw) as a function of T_1 using the signal equations for MPRAGE and FLASH sequences. We compute values of T_1 for several possible values of the ratio image. Then, given a value at the ratio image we can get a value for T_1. Since the sequence parameters are different for Siemens and GE, we compute separate lookup tables for each site. At first, the calculated R_1 values were off, and we believe this is due to differences in scaling between the T1w and PDw images. This can occur if each scan gets its own auto prescan or prescan normalize, or if receiver gains are different between scans. So, we also incorporate a scaling factor to account for not performing pre scan normalize properly. A single multiplicative factor was determined for each vendor by trial and error. We tried different scale values until our R_1 values were in the range of those reported in

the literature. The equations used for modelling MPRAGE signal and FLASH signal are described below:

MPRAGE: Signal intensity at the i[ith] readout pulse:

$$s_i \propto M_i \cdot \sin\theta = M_0 \cdot \sin\theta \cdot \left[\frac{(1-\delta)\cdot\left(1-\mu^{i-1}\right)}{1-\mu}\right] + \mu^{i-1}\cdot(1-\gamma) - \gamma\cdot\mu^{i-1}\cdot\frac{M_{eq}}{M_0} \quad (1)$$

And M_{eq} is

$$M_{eq} = \frac{1-\varphi+\frac{\varphi\cdot\cos\theta\cdot(1-\delta)\cdot(1-\mu^{N-1})}{1-\mu}+\varphi\cdot\cos\theta\cdot\mu^{N-1}+\rho\cdot\cos\alpha\cdot\cos^N\theta}{1-\rho\cdot\cos\alpha\cdot\cos^N\theta}\cdot M_0 \quad (2)$$

$\gamma = e^{-TI/T1}$

$\delta = e^{-\tau/T1}$

$\varphi = e^{-TD/T1}$

$\rho = e^{-TR/T1}$

θ is flip angle, α is inversion angle, TR is repetition time (TR = TI + N*ES + TD). N is total number of readout RF pulses. This equation is from Wang et al., (2014) and also appears in Protti et al., (2018). The MPRAGE sequence can be expressed equivalently in different ways (see Diechmann et al., 2000 or Bock et al., 2013).

FLASH: Signal intensity is given by

$$s_i \propto M_i \cdot \sin\theta = M_0 \cdot \sin\theta \cdot \frac{1-e^{-TR/T1}}{(1-\cos\theta)\cdot e^{-TR/T1}} \quad (3)$$

This is taken from Helms et al., (2008) and also described in detail in Wang et al., (2005). To see the effect of creating a ratio image, we can express the equations like Bock et al., (2013):

$$s_{MPRAGE} \propto M_{MPRAGE}|B_1^-| \cdot \sin(\theta_1|B_1^+|)$$

$$s_{FLASH} \propto M_{FLASH}|B_1^-| \cdot \sin(\theta_2|B_1^+|)$$

$$S_R \propto \frac{S_{MPRAGE}}{S_{FLASH}} \propto \frac{M_{MPRAGE}\cdot\sin(\theta_1|B_1^+|)}{M_{FLASH}\cdot\sin(\theta_2|B_1^+|)} \quad (4)$$

As mentioned in Bock et al., (2013), creating a ratio removes effects of B_1- but some B_1+ remains. We represent the scale factor here as g to account for differences in receiver/transmit gain and differences in prescan normalize.

$$g\frac{S_{MPRAGE}}{S_{FLASH}} (5)$$

This was coded in python and is available online (www.github.com/lyoganathan/MPRAGE_FLASH_LUT). We pass our ratio image and B_1+ into the lookup table and get a value of T_1 at each voxel using linear interpolation as implemented in scipy's griddata function.

4.7 To the Surface: HCP minimal processing pipeline using ciftify

We closely follow the HCP minimal preprocessing pipeline for surface creation and registration using ciftify (https://github.com/edickie/ciftify) (Dickie et al., 2019) and custom scripts for myelin mapping adapted from the HCP scripts (https://github.com/Washington-University/HCPpipelines) (Glasser et al., 2013). We use ciftify's recon-all pipeline which takes a Freesurfer output and transforms it into the CIFTI format used by the HCP. For myelin mapping, we use custom scripts that work with the output of ciftify_recon_all to generate myelin maps from our image data. We bring surface data into MATLAB using the gifti toolbox (https://www.artefact.tk/software/matlab/gifti/) and view our results with SurfStat (http://www.math.mcgill.ca/keith/surfstat/).

4.8 Statistical Correction for Multi-Site Data

We were also interested in comparing our quantitative MRI approach to statistical approaches common in multi-site studies, such as including site as a covariate in a general linear model. We fit a linear model to ratio image signal intensity at each vertex as a function of participant age, sex, site and B_1+:

$$Signal_{vp} = a_v + \beta_{1v}Age_p + \beta_{2v}Age_p^2 + \beta_{3v}Sex_p + \beta_{4v}Site_p + \beta_{5v}\Delta B_{1\,vp}^{+} + \varepsilon_{vp} \quad (6)$$

Signal at a vertex v is the value of the ratio image projected on to the surface for each participant p. α is the intercept and β's are coefficient for each term in the model. α and the β's are estimated using ordinary least squares (OLS). ε is the residuals, which is the observed signal minus the predicted signal from the model. To remove the linear effect of site and B_1+ from the signal, one can add the estimated coefficients multiplied by the variables of interest to the residuals and intercept. At a single vertex for each participant p:

$$Corrected\ Signal_{vp} = \hat{\alpha}_v + \varepsilon_{vp} + \hat{\beta}_{1v}Age_p + \hat{\beta}_{2v}Age_p^2 + \hat{\beta}_{3v}Sex_p \quad (7)$$

Alternatively, you can subtract the estimated coefficients multiplied by site and B1+ from the original data:

$$Corrected\ Signal_{vp} = Signal_{vp} - \hat{\beta}_{4v}Site_p - \hat{\beta}_{5v}\Delta B_{1\,vp}^{+}$$

We plot the average signal on an average surface as well as boxplots of average values across the cortex for each participant by site.

We apply another model called ComBat to our data, which originated as a method to correct batch effects in microarray genome sequencing, but has also been

used to correct multi-site DTI and cortical thickness data. You can further correct the ComBat adjusted data for B_1+ using a linear model as above, but this analysis is not presented here. ComBat is an extension of location scale models, and uses empirical bayes to improve parameter estimates by calculating a prior for the location and scale parameters from the data. This is the model:

$$y_{ijv} = \alpha_v + X_{ij}\beta_v + \gamma_{iv} + \delta_{iv}\varepsilon_{ijv} \quad (8)$$

where y is signal at site i scan j and voxel v. α is the intercept, X is the matrix of variables of interest, and β is a vector of estimates of coefficients for each variable. γ is the location parameter and δ is the scale parameter.

Step 1: α and β are estimated with OLS on the data with this constraint on γ:

$$\sum_i n_i \hat{\gamma}_{iv} = 0$$

Where n_i is number of subjects in each site i. Pooled variance is also calculated and used in step 2:

$$\hat{\sigma}_v^2 = \frac{1}{N}\sum_{ij}\left(Y_{ijv} - \hat{\alpha}_v - X\hat{\beta}_v - \hat{\gamma}_{iv}\right)^2$$

Step 2: Standardize data:

$$Z_{ijv} = \frac{y_{ijv} - \hat{\alpha}_v - X_{ij}\hat{\beta}_v}{\hat{\sigma}_v}$$

This above equation is basically (data – mean) / pooled variance.

Step 3: Calculating γ and δ

We assume that

$$Z_{ijv} \sim N(\gamma_{iv}, \delta_{iv}^2)$$

We can estimate the location parameter on standardized data with OLS. The standardized data is adjusted for covariates, and now site is the only independent variable:

$$Z_{ijv} = \gamma_{iv} + \varepsilon_{ijv}$$

The location parameter is the mean of each group. The scale parameter is the standard deviation of each group.

Step 4: Empirical Bayes

The estimate for location and scale can be improved by using prior distributions calculated from the data. See supplemental in Johnson, Li, & Rabinovic (2007) for full derivation and explanation.

Step 5: Adjust Data

$$y_{ijv}^{ComBat} = \frac{y_{ijv} - \hat{\alpha}_v - X_{ij}\hat{\beta}_v - \gamma_{iv}^*}{\delta_{iv}^*} + \hat{\alpha}_v + X_{ij}\hat{\beta}_v \quad (9)$$

After comparing the different methods in surface space, we also wanted to investigate some whole brain multi-site data that has been adjusted using ComBat. So we also run ComBat in volume space and look at correlations between volume and signal in the deep white matter.

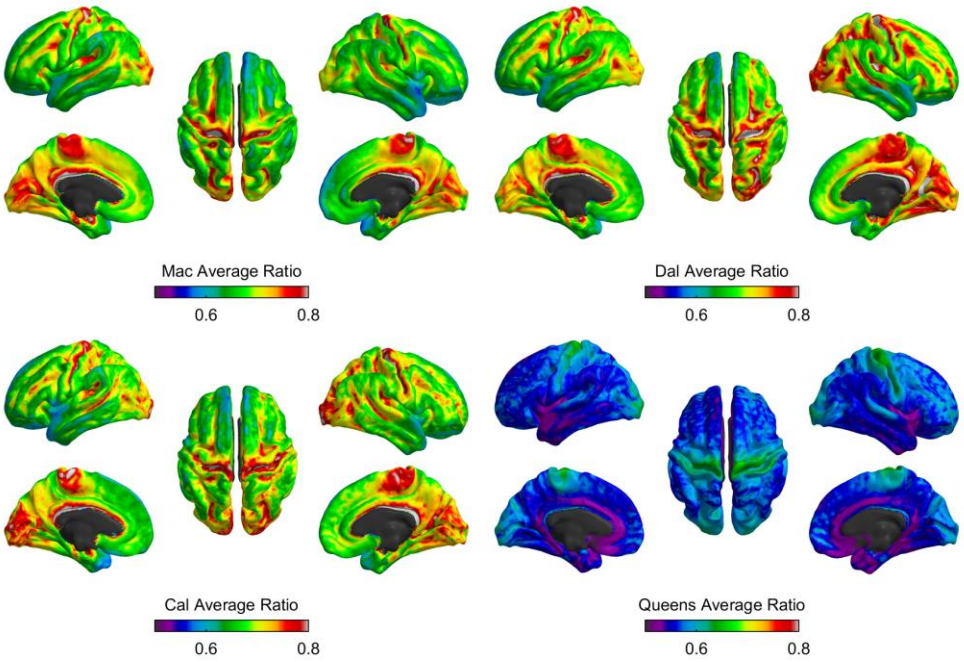
To run ComBat on the ratio image in volume space and also in surface space, implementations are available online in both MATLAB and R ((https://github.com/Jfortin1/ComBatHarmonization).

5. Multi-Site Results:

The figures below show average surface maps for all the healthy controls at baseline for each site. The maps are created by projecting intensity values in the cortex onto the surface. We show the ratio image, B_1+ maps, R_1 maps & statistically corrected maps. The boxplots represent average intensity across the cortex for all participants at each site.

5.1 Ratio Image Signal Intensity:

First, we show the bias corrected ratio image created by dividing a high contrast T1w MPRAGE image by a PDw FLASH image. Different pulse sequences were used at the Siemens and GE scanners and so a difference between vendors is expected. Across the three GE sites, the ratio image signal intensity is comparable. At the Siemens site, the ratio image intensity is low enough that we cannot simply pool this data together, demonstrating the need for a correction for site effects. In the boxplot, we can also see that the within subject variance is lower in the Siemens site compared to the GE sites.



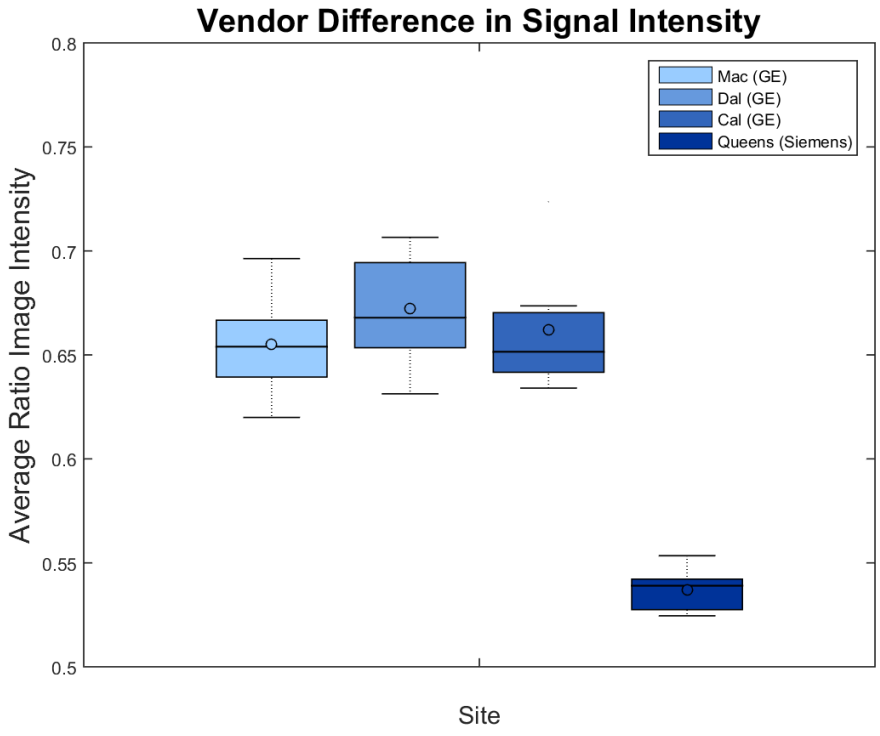
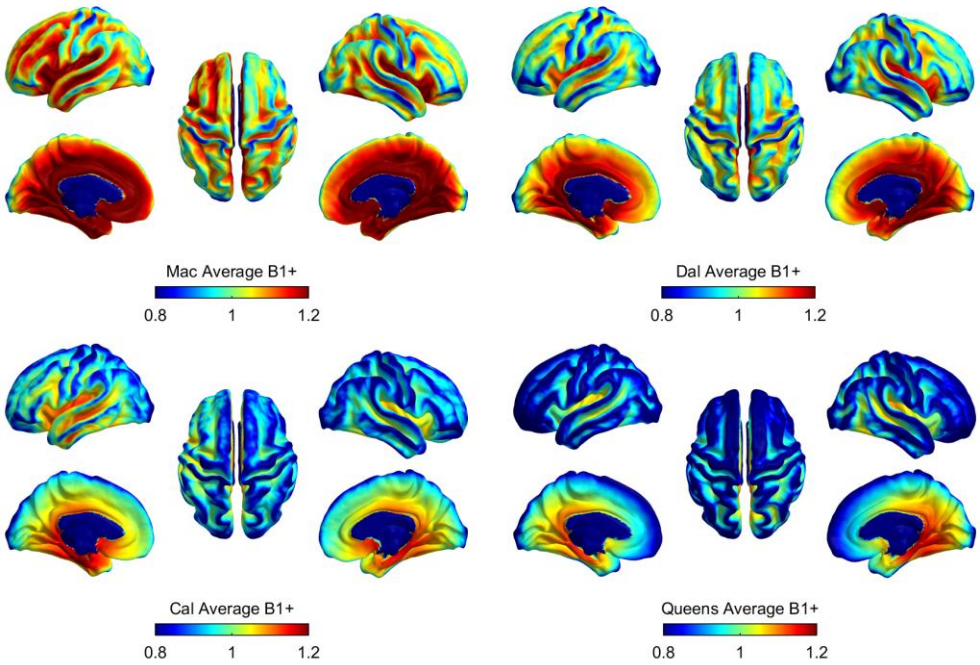


Figure 10: Ratio (T1w/PDw) Signal intensity by site. Signal intensity is comparable between the three GE sites, but Siemens is lower than GE. This is expected because we used different pulse sequences for GE & Siemens.

5.2 B_1+ Mapping

For GE, the Bloch-Seigert shift B_1+ mapping returns values of 10*flip angle. In this case, the prescribed flip angle was 15, so we divide the B_1+ image by 150 for each GE site to get values of B_1+. The double angle method calculates B_1+ using two spin echo images. We can see that across GE sites there is difference in B_1+ mapping. We also note differences between double angle and Bloch-Siegert shift methods. Since this difference in B_1+ mapping will affect both R_1 calculations and including B_1+ as a covariate in a linear model, we decided to mean center the maps at each site to 1 for R_1 calculation, and to 0 for linear modelling. To center the B_1+ maps at 1, we subtract each individuals B_1+ map from the site mean (center at 0) then add 1. The reason for mean centering to 0 for linear modelling is so that we can subtract the effects of B_1+ and the corrected signal will be on a similar scale to uncorrected ratio image signal intensity for comparison purposes.



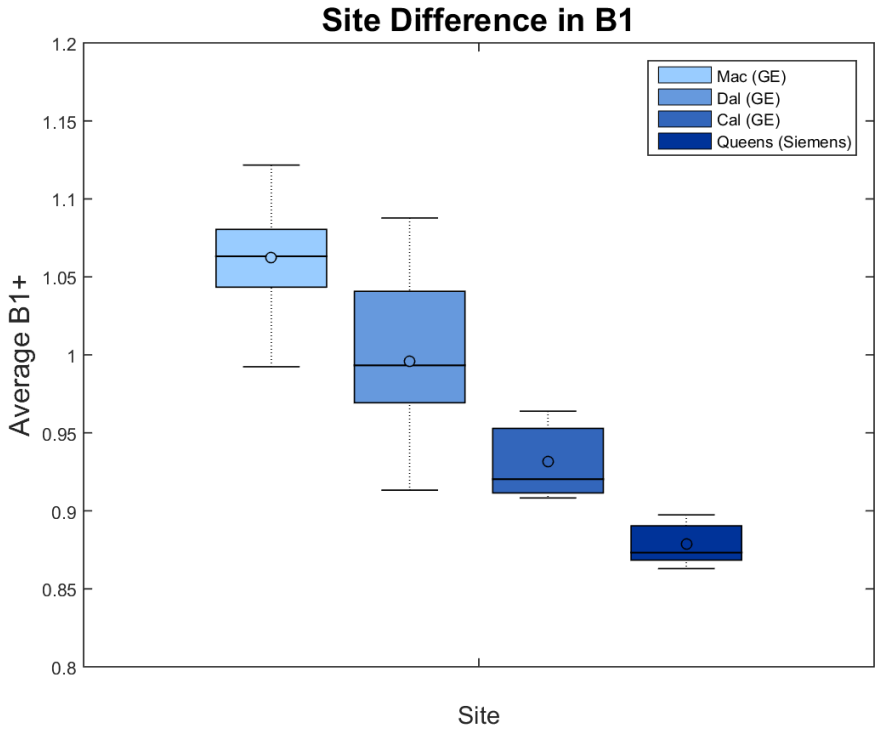


Figure 11: B1+ maps by site. Even though the three GE sites used the exact same B1+ mapping protocol, there were significant differences across sites. There also seems to be a difference due to B1+ mapping method as the Siemens maps (double angle) are lower than the GE maps (Bloch-Siegert shift)

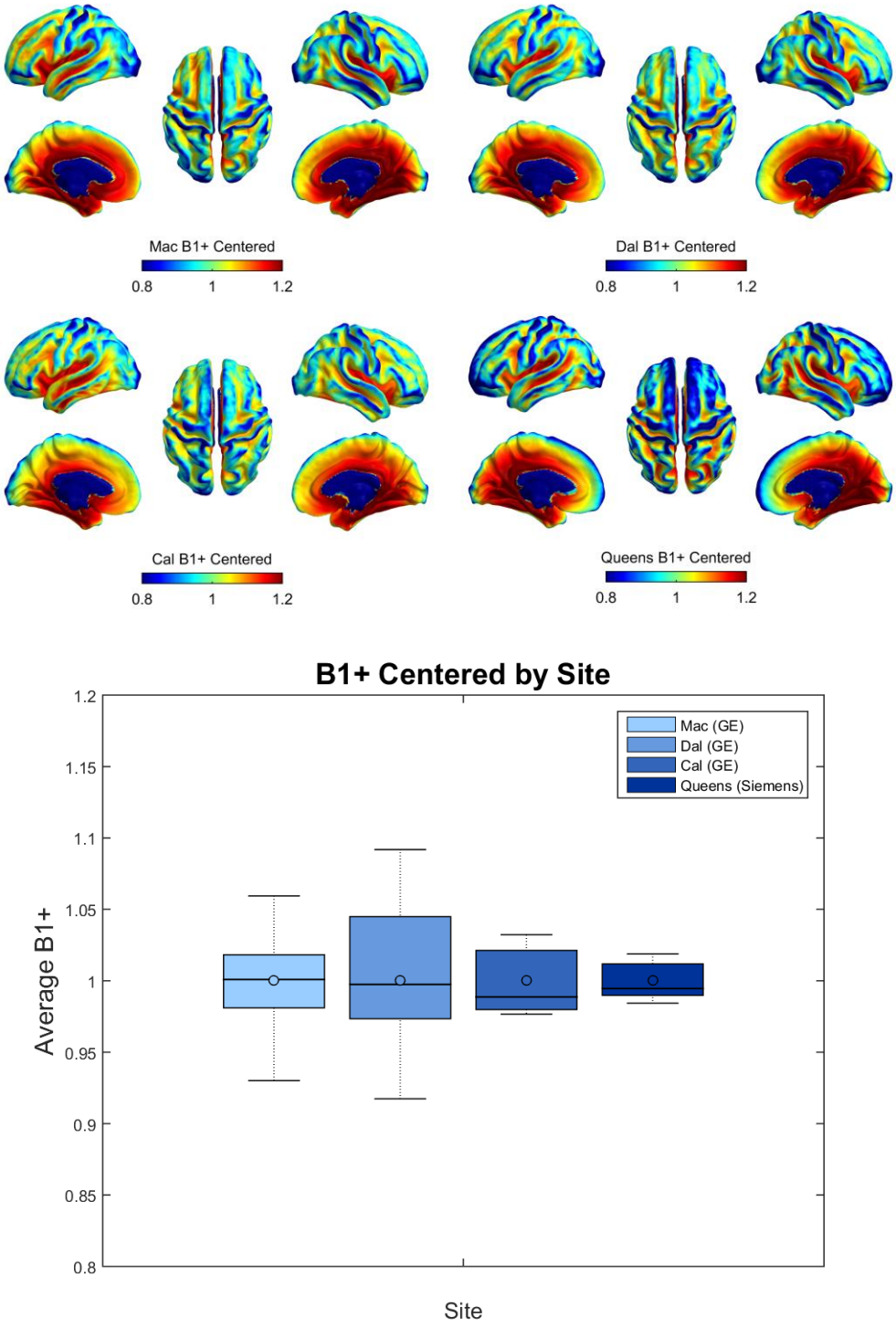
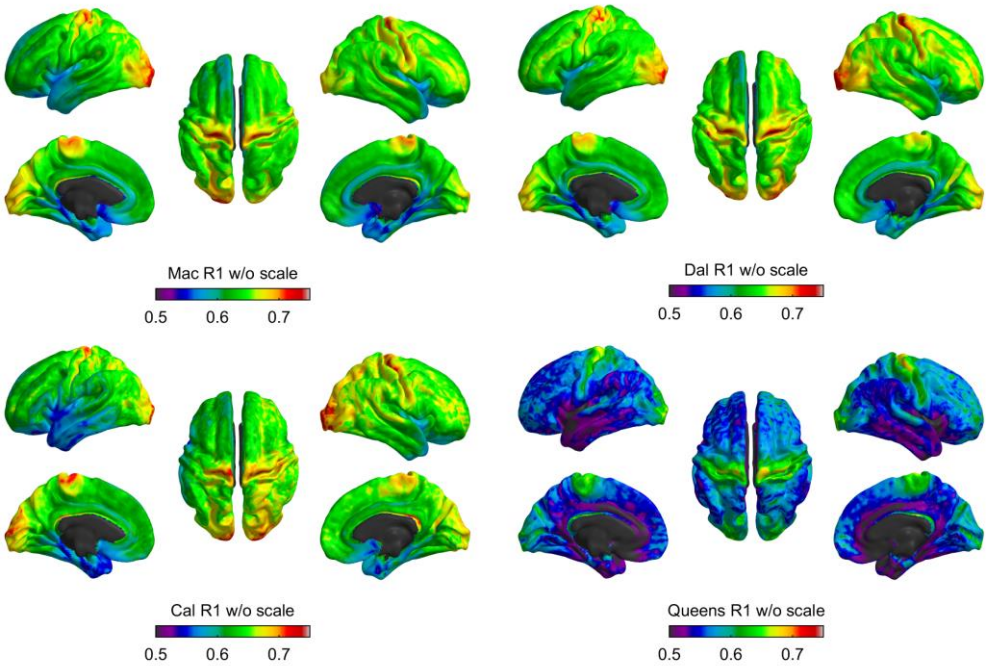


Figure 12 B1+ maps centered at 1 by site. This is for R1 mapping. Without centering B1+ maps, estimated R1 values will be different across sites.

5.3 R_1 with B_1+ correction:

We estimate R_1 with and without a vendor specific scale factor. We notice that without a scale factor, our R_1 method does not address the difference between vendors. We also see that our R_1 values are lower than those reported previously in the literature. By incorporating a scaling term for vendor, we can get the R_1 values between vendors closer to each other, and closer to values reported in the literature. However, we notice that the within subject variance in the Siemens site now seems greater than the GE sites. This is also evident on the cortical surface plots as the Siemens site seems to show a greater range of values across the cortex.



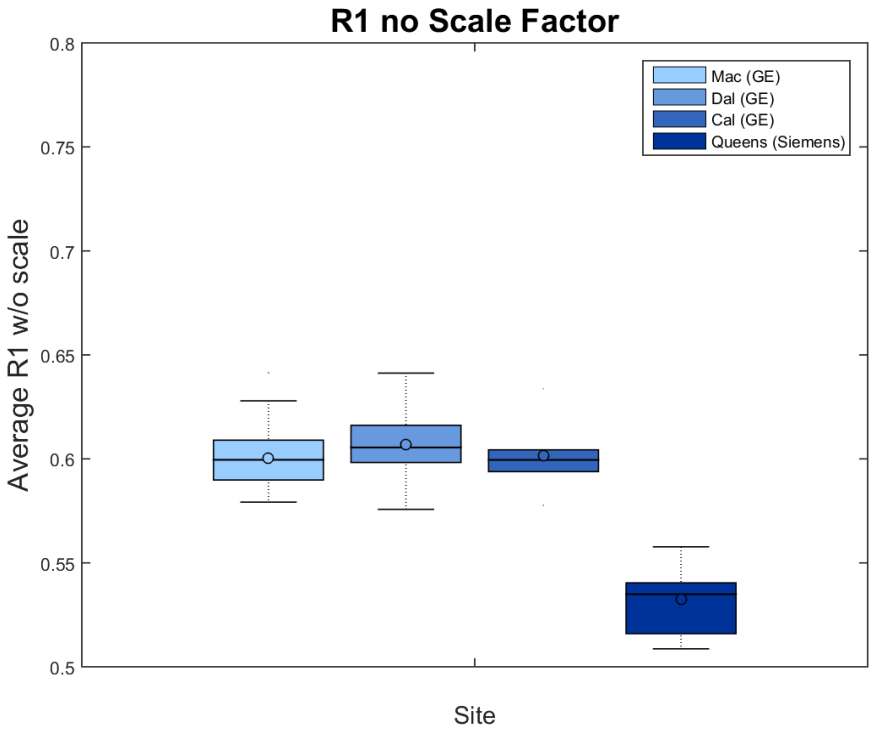


Figure 13: R1 estimated without a vendor-specific scale factor. This is not much different from original ratio image signal intensity. This suggests that there is an error in our R1 method, which we suspect is coming from differences in receiver gain.

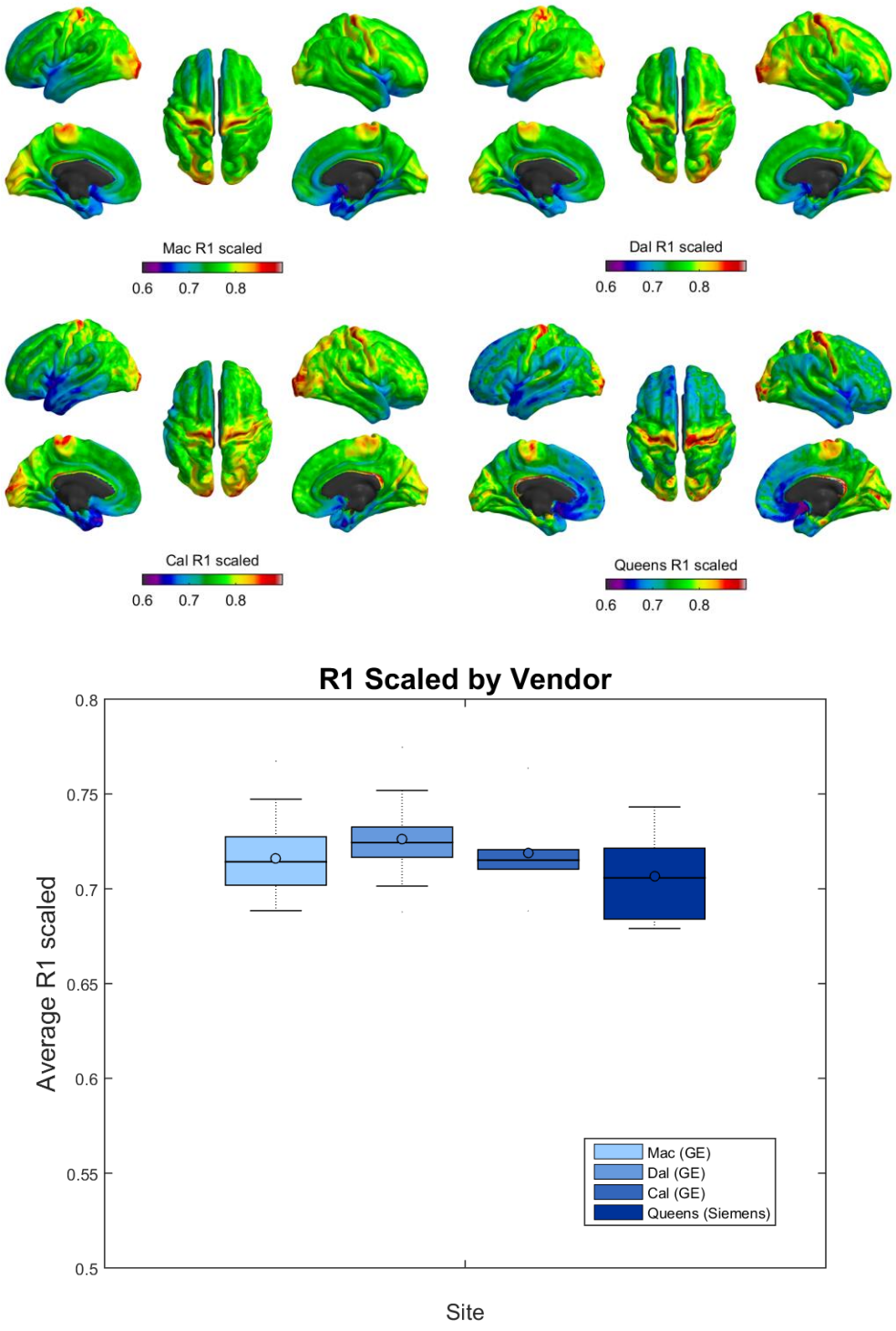
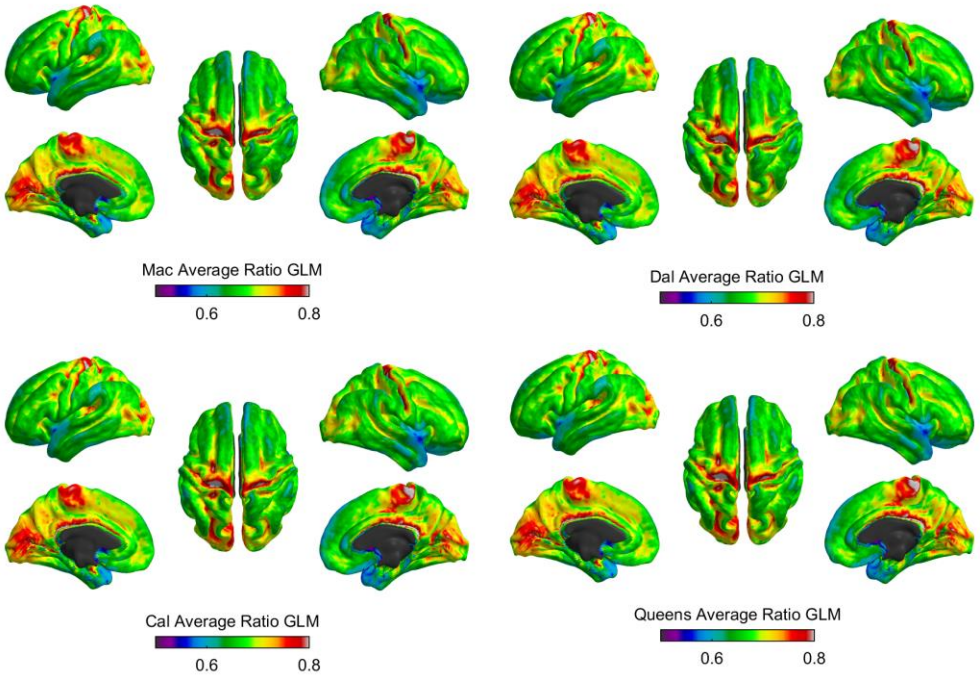


Figure 14: R1 estimated with a vendor specific scale factor. The scale factor was estimated by trial and error until the Siemens site was close to the other GE sites. We multiplied the GE ratio data by 1.35 and the Siemens data by 1.2

5.4 Adding Site as a covariate in a linear model:

Here we use a vertex-wise linear model with ratio image signal as the dependent variable and age, sex, B_1+ and site as independent variables. We remove the effects of site and B_1+ at each vertex. Although the average myelin maps look comparable across site, a boxplot of the average intensity values for each participant across the cortex show that within subject variation is lower in the Siemens site. This approach corrects for mean differences across sites, but does not correct differences in variance. However, it may also be the case that the difference in variance is coming from actual differences in participants as different people were imaged at each site.



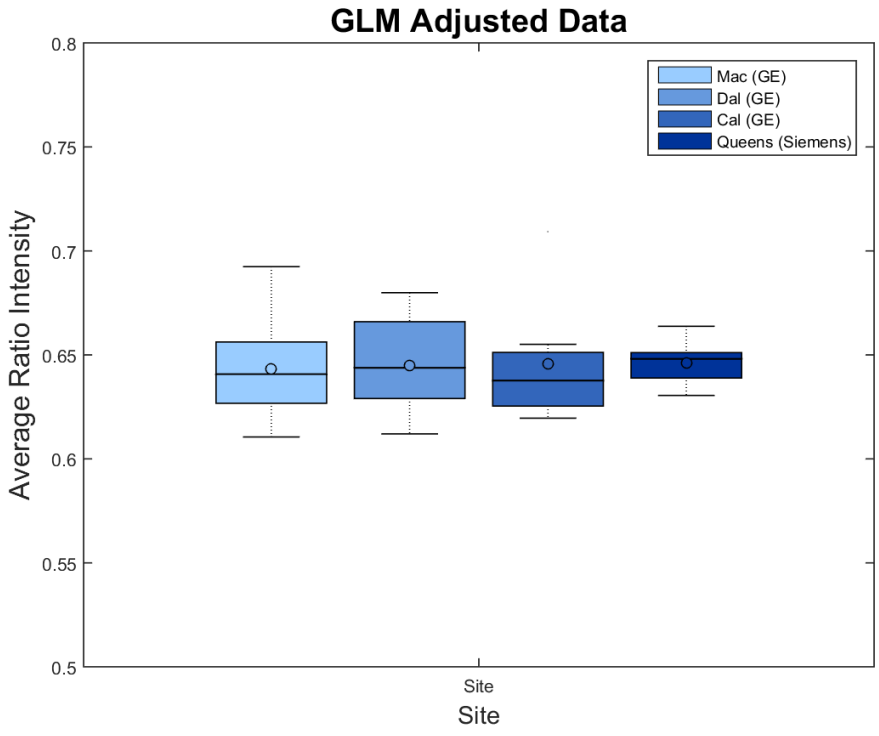
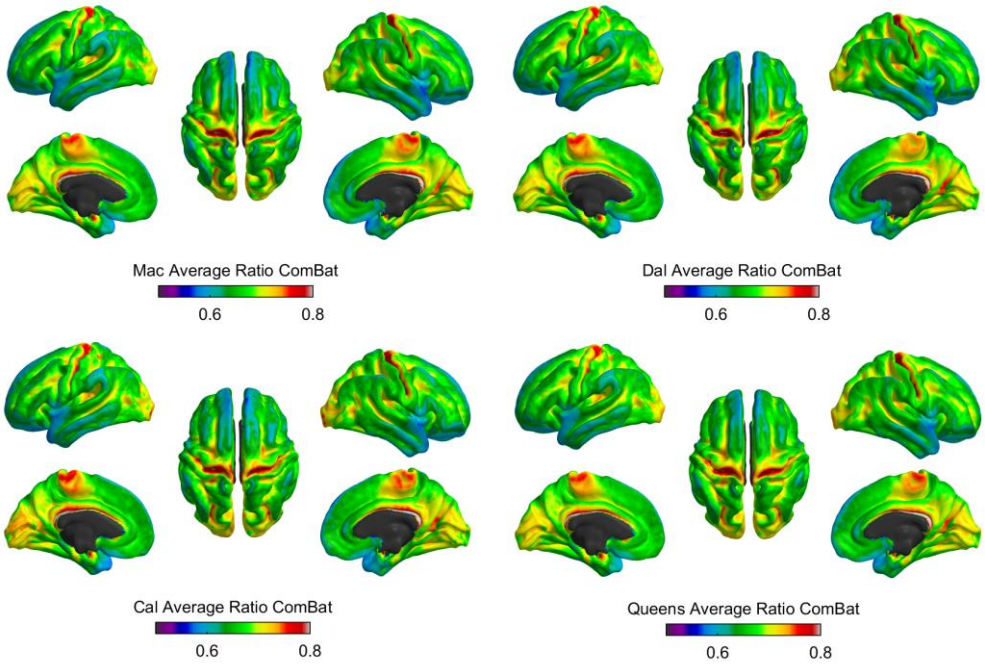


Figure 15: Ratio image adjusted for site and B1+ using a vertex-wise GLM

5.5 Using ComBat method to adjust data:

When we adjust the data using the ComBat method, the average myelin maps look comparable and distribution of values also look comparable across sites. However, since we have no ground truth, there is no way of determining which method is best. Variation in the data could be scanner artifacts or real variation in participants. It could be the case that ComBat is removing real subject to subject variation from the data.



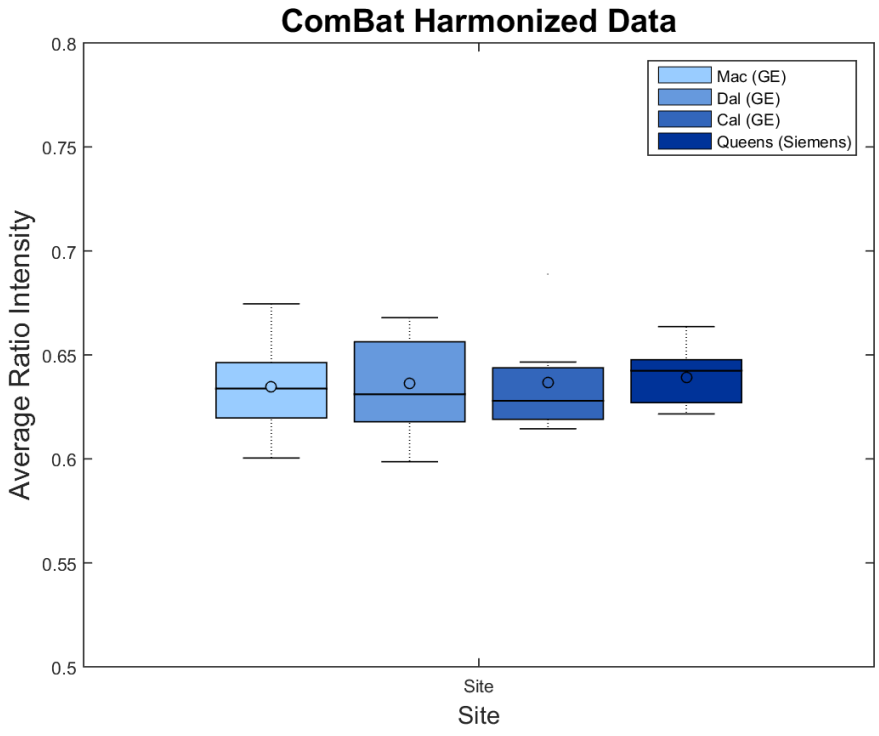


Figure 16: Ratio image adjusted using the ComBat method

6. Relationship between Volume and Signal intensity in White Matter

After comparing multi-site correction approaches, we also compare signal intensity and volume in the deep white matter (WM) across sites. Volumes were obtained from freesurfer's aseg and wmparc segmentation. Signal intensity was corrected using the ComBat method. We see that volume does not have a site or vendor effect like signal.

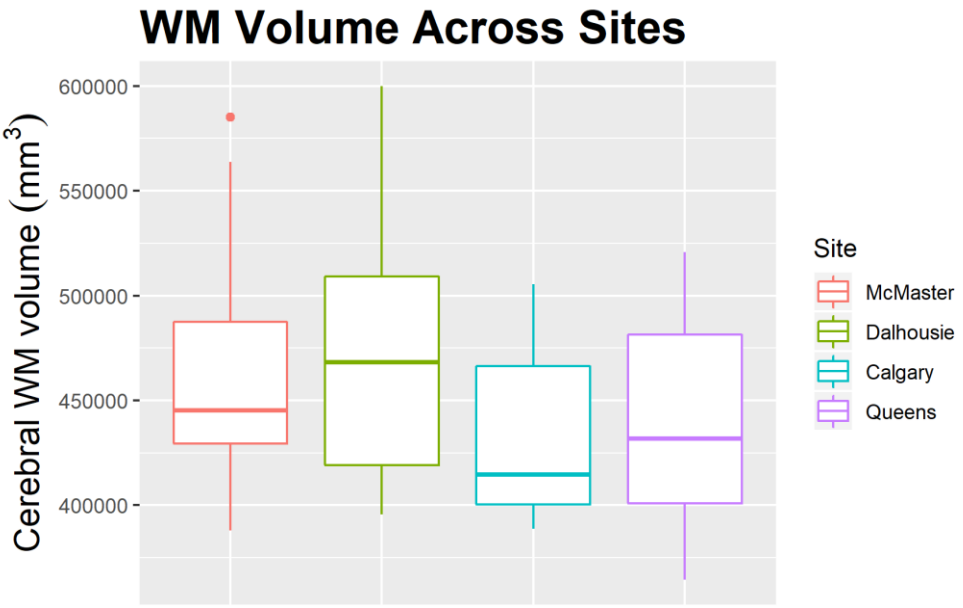


Figure 17: Deep WM volume across sites.

Although both signal intensity and volume have been used to study myelin, few studies look at both metrics. We take correlations between WM Volume and WM signal (ComBat Harmonized) stratified by sex, because volume has a stronger sex effect than signal which might affect correlation values. WM signal is calculated by taking the average signal in the entire deep WM segmentation. Volume is number of voxels in the entire deep WM segmentation.

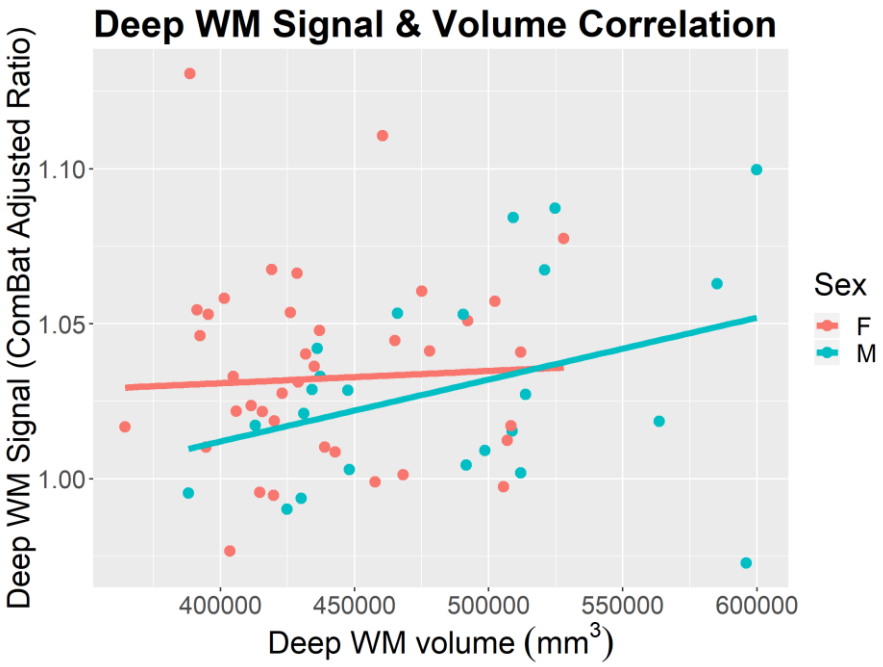


Figure 18: Correlation between deep WM signal and volume in males and females

The correlation coefficient r in males is 0.4 and in females is 0.003. We can look at age trajectories of volume and signal to see where this difference is coming from. Below is a plot of WM signal harmonized with ComBat pooled across the 4 sites fit with a quadratic model. We see that relationship between age and females is obscured by some outliers with high signal in the young age range.

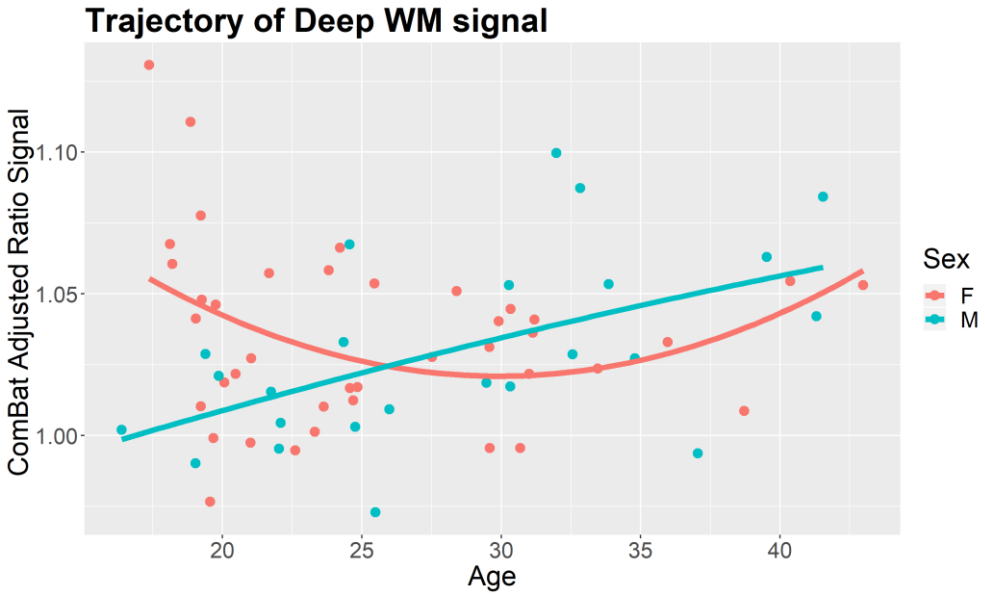


Figure 19: Age trajectory of deep WM signal in males and females. Males seem to increase and begin to plateau whereas the trajectory in females seems to follow a U shaped curve

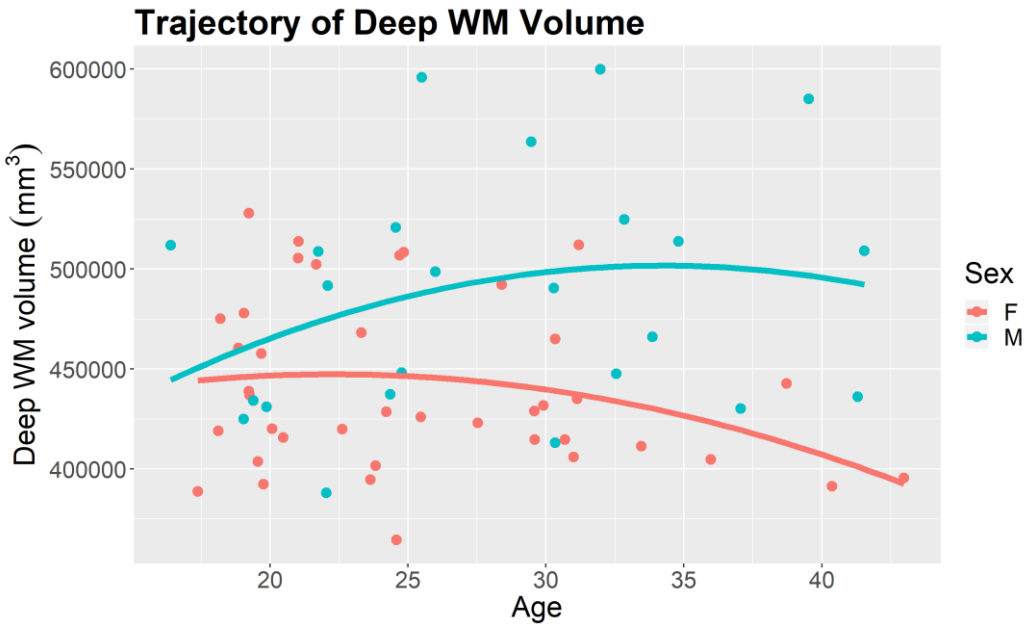


Figure 20: Age trajectory of deep WM volume in males and females. Both curves follow an inverted quadratic trajectory similar to that of signal and myelin

However, looking at the age trajectory of deep white matter volume, we see that there are not really high values in the low age range females. This suggests that the high signal intensity seen in females may be some MRI artifact affecting signal rather than actually reflecting higher levels of myelin. Therefore, the low correlation in females might be due to an issue with MRI signal in young females of this sample. Both volume and signal might be reflecting changes in myelin, although both are indirect measures of myelin and might be sensitive to different sources of error or bias.

7. Discussion/Limitations:

We used structural MRI at four sites to investigate brain myelination and found significant intervendor differences in signal intensity of the scans, but no differences in the volume of brain structures across sites. Although our data show no significant differences in volume, multi-site differences in volume have been reported in other studies. In the ADNI study which was performed across 98 scanners with different vendors and field strengths, they found significant differences in volume measurements (Kruggel et al., 2010). In this thesis, we compared 3 approaches to correct for multi-site differences observed in signal. We tried to use quantitative MRI by calculating R1, adjusting our data by using site as a covariate in a linear model and also using the ComBat method.

7.1 R_1 Calculation and Scaling:

Our R1 method did not work. Our calculated R_1 values were lower those reported in the literature and also different across scanner vendors. We suspected this could be due to a mistake with image scaling differences between the T1w and PDw image,

which would throw off the R_1 calculation. The equations presented in the methods section are simplified. We do not show some things like receiver gain or T_2^* decay which also affect signal intensity. T_2^* effects should be the same in both scans, and we assume that gain effects are also the same between scans. The ratio image division should remove these effects. In equations 4 and 5, our model of MRI signal expects specific values for MPRAGE and FLASH signal intensity based on our sequence parameters. However, if different sequences get scaled differently due to differences in gain (usually the default behaviour of the scanner), the R_1 lookup table will not provide accurate mapping between ratio image values and R_1. If not set up properly in the protocol beforehand, the scanner automatically sets receiver gain for each sequence as part of the auto-prescan or prescan normalize step. This can be avoided by applying auto-prescan to the first scan only (usually the one with highest intensity which is usually the T1w scan), then using manual prescan or turning prescan normalize off for subsequent scans. This should ensure that gains are consistent between scans.

$$S_R \propto \frac{g_1 S_{MPRAGE}}{g_2 S_{FLASH}} = g \frac{S_{MPRAGE}}{S_{FLASH}}$$

We represent the automatic scaling applied to each image as g_1 and g_2. Assuming g_1 and g_2 are different, we evaluated whether we could correct this error in our R_1 mapping using a single scale factor g to make R_1 values at all sites match each other and the range of reported values in the literature. We tried several values of g until our R_1 values were in a reasonable range. Another method to determine a value for g could be by measuring R_1 using inversion recovery and comparing it to R_1 values given by our method. Using a single scale factor assumes there is no non-linear scaling of intensity. If there was nonlinear scaling going on, we would have to figure out how scaling was applied to each image and we could not correct the differences in scaling with a single term. If a method such as histogram matching was used to scale the images (Nyúl & Udupa, 1999), then this single scale factor approach would not be valid.

Another problem is that we are not entirely sure if the difference in calculated R_1 is due to scaling. It could also be a problem with our R_1 equations and modelling. One way to verify if our equations work as intended would be to test several flip angles and inversion times to make sure that our R_1 estimates are accurate for a range of sequence parameters. This would also help in quantifying the error in our measurement method. Also, our R_1 method has not been validated against inversion recovery measurements which is the gold standard for measuring R_1. Typically, when testing a new R_1 method an inversion recovery measurement is also performed in order to evaluate the accuracy and precision of the R_1 method. For example, in the MP2RAGE paper, to demonstrate accuracy of their method, Marques et al. (2010) perform MP2RAGE with many different inversion times. They also perform single slice inversion recovery. They show agreement between both methods in a head phantom at 3T and 7T. Standardized protocols have also been developed to deal with multi-site MRI problems such as the multi parametric mapping (MPM) approach that have been

validated at multiple sites (Weiskopf et al., 2013). MPM makes quantitative maps of R_1, R_2^*, PD^* and MTR that are consistent across sites and vendors. More established methods which have been validated against inversion recovery measurements might have given more consistent results between sites. Lee et al., (2019) showed good agreement between vendors using the variable flip angle approach for T_1 mapping. They also discuss the importance of establishing good intra and inter scanner/vendor reliability for quantitative metrics during the piloting phase of a study, which we also recommend.

Another issue which affected R_1 is the method used for B_1+ mapping. The Bloch-Seigert shift method seemed to vary considerably within GE sites, even though the prescribed protocol was the exact same. A study comparing the double angle method, actual flip angle and Bloch-Siegert shift showed that all methods performed similarly, (Boudreau et al., 2017), although this comparison was made on the same scanner. There are currently no guidelines on the best way to map B_1+ across multiple sites. Future studies should also try to establish reliability between B_1+ mapping methods during piloting. It also appears that the B_1+ correction is "overcorrecting" in some areas. Our B_1+ corrected R_1 maps show relatively lower values in the posterior cingulate cortex (PCC) compared to other published R_1 maps and compared to myelin maps made with other imaging modalities such as T1w/T2w ratio (Shams, Norris, & Marques, 2019, supplementary figure 2 & 3).

7.2 Site as a Covariate in a Linear Model:

We also tried to correct our data statistically using linear modelling with site as a covariate as this is a common technique in the literature along with variants such as using site as a random effect in a mixed effects model (Fennema-Notestine et al., 2007). We found that this may not always be enough to correct for scanner differences. It corrected for mean differences across site but it did not correct for differences in variance between sites. Other studies have used site as a covariate when analysing large multi-site public data sets. Kwon et al. (2018) looked at myelin and developmental trajectories using site as a covariate in the NCANDA dataset. They looked across adolescence to adulthood. They did not indicate if site effects were present or plot their data by site, however they did use normalized myelin content using the T1w/T2w, which may be affected differently by multi-site scanner variation compared to R_1. Similarly, Hibar et al., (2018) looked at cortical grey matter abnormalities in bipolar disorder and used site as a random effect in a mixed effects model. They also did not discuss effect of site or plot data by site. Although a convenient method, our results show it may not have been enough to fully correct for site differences. We believe a more sophisticated method such as ComBat is better suited to correct multi-site differences statistically.

7.3 Combat Method:

We found that the ComBat correction gave the best similarity between sites. It corrected for mean differences in intensity as well as differences in within subject

variance. Fortin et al., (2017) compared combat to 4 other methods and showed it performed the best in terms of preserving biological variability while removing unwanted scanner variability. It also increased power of subsequent statistical analyses. This was looking at fractional anisotropy, in DTI data. Combat has also been used to correct cortical thickness data with similar results (Fortin et al., 2018). It has also been used to correct site effects in connectivity and network metrics derived from fMRI data (Yu et al., 2018). In order to test if ComBat performs better than quantitative metrics at correcting for site effects, we would need to image the same people at multiple sites.

7.4 Sampling Issues:

Sample size affects how well statistical correction methods work. Larger sample sizes ensure the model estimates site effects accurately. For example, to estimate the effect of B_1+ B_1+, estimating the relationship between each vertex in the ratio image and each vertex in the B_1+ map, having a large sample makes the statistical correction better. This is an advantage of R_1 and other quantitative MRI metrics over statistical methods. Since R_1 is calculated at the imaging level, sample size and unbalanced groups have no effect on how well R_1 works.

In our dataset, there is a difference in the number of males and females at each site. At the Queens site, we have 2 males and 10 females, whereas at the Calgary site we have 7 males and 1 female. We did not thoroughly investigate any effects caused by the difference in sampling at each site. Other studies have looked at how the ratio of males and females in a sample affect multi-site results and how adding site as a covariate might interact with this difference in samples (Takao, Hayashi, & Ohtomo, 2014). However, there is not much literature on the increase in signal intensity we see in young females. We believe that there are a few possibilities for the outliers seen in our female sample. Previous work has discussed the role of sex hormones on myelin content, even showing transient changes in myelin across the menstrual cycle (Barth et al., 2016) Another possibility is that the MRI signal is being increased by something other than myelin in young females such as medication like iron supplements. As opposed to WM signal, WM volume seems to follow an inverted quadratic age trajectory in females. A future study could use the same subjects across multiple sites and also ensure samples are balanced to better characterize the relationship between different MRI metrics. This design would also be useful to evaluate which correction method achieves greatest similarity between sites.

7.5 Signal Intensity and Volume:

Both white matter volume and signal intensity are used as proxies for myelin. Studies often refer to changes in white matter volume as changes in myelin, or changes in white matter signal as changes in myelin, but both are indirect measures of myelin. We found that signal intensity was different across vendors whereas volume was not, suggesting that these metrics may be measuring different things. The correlation in

males and the difference in females suggests that although both measures may be sensitive to myelin, they are also affected by different sources of noise.

Studies on white matter volume generally look at deep white matter volume, although cortical white matter volume has also been studied (Bartzokis et al. 2009, Rowely 2015, Tishler 2019). Volume measurements are affected by the contrast in an MRI image. The boundaries for different tissue classes (CSF, GM, WM) are defined based on gradients of signal intensity. Volume measurements are normally acquired by segmenting a T1w image corrected for intensity nonuniformity by using a method such as N3. However, using different MRI contrasts could also give different estimates for volume of brain regions. Haast et al. (2018) showed that correcting MP2RAGE data for B_1+ gave significantly different results for estimates of cortical thickness in some brain areas using Freesurfer. Helms et al., (2009) obtained better segmentations and volume estimates from magnetization transfer maps. Another possible future study could look at which image (T1w/PDW vs R_1 vs T1w with N3 correction) gives better estimates for volumes across multiple sites.

We have only discussed site differences in structural MRI and evaluated just 3 potential correction methods. However, site effects have been demonstrated in other MRI modalities and several other methods exist to correct multi-site data. For future studies we recommend establishing reliability and repeatability during the piloting phase. This is especially important for quantitative imaging metrics as there are many variables that need to be taken to account. Also, we recommended building on previous research and using methods which have more support. For example, open source software such as qMRLab (https://qmrlab.org/) has code to calculate quantitative metrics for several popular quantitative MRI sequences, which would be much easier than programming custom calculations. Finally, we also recommend combining and comparing different metrics that can be obtained from the same MRI scan to gain a better understanding of changes in the brain.

References:

Bartzokis, G. (2005). Brain myelination in prevalent neuropsychiatric developmental disorders: Primary and comorbid addiction. Adolescent Psychiatry, 29, 55.

Bartzokis, G., Lu, P. H., Stewart, S. B., Oluwadara, B., Lucas, A. J., Pantages, J., et al. (2009). In vivo evidence of differential impact of typical and atypical antipsychotics on intracortical myelin in adults with schizophrenia. Schizophrenia research, 113(2-3), 322-331.

Barth, C., Steele, C. J., Mueller, K., Rekkas, V. P., Arélin, K., et. al (2016). In-vivo dynamics of the human hippocampus across the menstrual cycle. Scientific reports, 6, 32833.

Bock, N. A., Hashim, E., Janik, R., Konyer, N. B., Weiss, M., Stanisz, G. J., Turner R., & Geyer, S. (2013). Optimizing T1-weighted imaging of cortical myelin content at 3.0 T. *Neuroimage*, *65*, 1-12.

Boudreau, M., Tardif, C. L., Stikov, N., Sled, J. G., Lee, W., & Pike, G. B. (2017). B1 mapping for bias‐correction in quantitative T1 imaging of the brain at 3T using standard pulse sequences. Journal of Magnetic Resonance Imaging, 46(6), 1673-1682.

Chenevert, T. L., Malyarenko, D. I., Newitt, D., Li, X., Jayatilake, M., Tudorica, A., et al. (2014). Errors in quantitative image analysis due to platform-dependent image scaling. Translational oncology, 7(1), 65.

Deichmann, R., Good, C. D., Josephs, O., Ashburner, J., & Turner, R. (2000). Optimization of 3-D MP-RAGE sequences for structural brain imaging. *Neuroimage*, *12*(1), 112-127.

Dickie, E. W., Anticevic, A., Smith, D. E., Coalson, T. S., Manogaran, M., Calarco, N., et al. (2019). ciftify: A framework for surface-based analysis of legacy MR acquisitions. NeuroImage.

Edwards, L. J., Kirilina, E., Mohammadi, S., & Weiskopf, N. (2018). Microstructural imaging of human neocortex in vivo. Neuroimage, 182, 184-206.

Eickhoff, S., Walters, N. B., Schleicher, A., Kril, J., Egan, G. F., Zilles, K., Watson J. D. G., & Amunts, K. (2005). High-resolution MRI reflects myeloarchitecture and cytoarchitecture of human cerebral cortex. Human brain mapping, 24(3), 206-215.

Fennema-Notestine, C., Gamst, A. C., Quinn, B. T., Pacheco, J., Jernigan, T. L., Thal, L., et al. (2007). Feasibility of multi-site clinical structural neuroimaging studies of aging using legacy data. Neuroinformatics, 5(4), 235-245.

Fortin, J. P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., et al. (2017). Harmonization of multi-site diffusion tensor imaging data. Neuroimage, 161, 149-170.

Fukunaga, M., Li, T. Q., van Gelderen, P., de Zwart, J. A., Shmueli, K., Yao, B., et. al. (2010). Layer-specific variation of iron content in cerebral cortex as a source of MRI contrast. Proceedings of the National Academy of Sciences, 107(8), 3834-3839.

Fukutomi, H., Glasser, M. F., Zhang, H., Autio, J. A., Coalson, T. S., Okada, T., Togashi K., Van Essen D. C., & Hayashi, T. (2018). Neurite imaging reveals microstructural variations in human cerebral cortical gray matter. Neuroimage, 182, 488-499.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., et al. (2013). The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage, 80, 105-124.

Harkins, K. D., Xu, J., Dula, A. N., Li, K., Valentine, W. M., Gochberg, D. F., Gore, C. J., & Does, M. D. (2016). The microstructural correlates of T1 in white matter. Magnetic resonance in medicine, 75(3), 1341-1345.

Haast, R. A., Ivanov, D., & Uludağ, K. (2018). The impact of correction on MP2RAGE cortical T 1 and apparent cortical thickness at 7 T. Human brain mapping, 39(6), 2412-2425.

Helms, G., Dathe, H., & Dechent, P. (2008). Quantitative FLASH MRI at 3T using a rational approximation of the Ernst equation. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 59(3), 667-672.

Helms, G., Draganski, B., Frackowiak, R., Ashburner, J., & Weiskopf, N. (2009). Improved segmentation of deep brain grey matter structures using magnetization transfer (MT) parameter maps. Neuroimage, 47(1), 194-198.

Hibar, D. P., Westlye, L. T., Doan, N. T., Jahanshad, N., Cheung, J. W., Ching, C. R., ... & Krämer, B. (2018). Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA Bipolar Disorder Working Group. Molecular psychiatry, 23(4), 932.

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics, 8(1), 118-127.

Kruggel, F., Turner, J., Muftuler, L. T., & Alzheimer's Disease Neuroimaging Initiative. (2010). Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. Neuroimage, 49(3), 2123-2133.

Kwon, D., Pfefferbaum, A., Sullivan, E. V., & Pohl, K. M. (2018). Regional growth trajectories of cortical myelination in adolescents and young adults: longitudinal validation and functional correlates. Brain imaging and behavior, 1-25.

Lee, Y., Callaghan, M. F., Acosta-Cabronero, J., Lutti, A., & Nagy, Z. (2019). Establishing intra-and inter-vendor reproducibility of T1 relaxation time measurements with 3T MRI. Magnetic resonance in medicine, 81(1), 454-465.

Leuze, C. W., Anwander, A., Bazin, P. L., Dhital, B., Stüber, C., Reimann, K., Geyer, S., & Turner, R. (2012). Layer-specific intracortical connectivity revealed with diffusion MRI. Cerebral Cortex, 24(2), 328-339.

Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P. F., & Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage, 49(2), 1271-1281.

Narvacan, K., Treit, S., Camicioli, R., Martin, W., & Beaulieu, C. (2017). Evolution of deep gray matter volume across the human lifespan. Human brain mapping, 38(8), 3771-3790.

Naiji, L., Yu, H., Tian, C., Douglas, D. G., Yinglin, X., Julia, Y. L., & Xin M.T. (2013). Power analysis for cross-sectional and longitudinal study designs. Shanghai archives of psychiatry, 25(4), 259.

Nieuwenhuys, R. (2013). The myeloarchitectonic studies on the human cerebral cortex of the Vogt–Vogt school, and their significance for the interpretation of functional neuroimaging data. Brain Structure and Function, 218(2), 303-352.

Nyúl, L. G., & Udupa, J. K. (1999). On standardizing the MR image intensity scale. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 42(6), 1072-1081.

Protti, A., Jones, K. L., Bonal, D. M., Qin, L., Politi, L. S., Kravets, S., Nguyen, Q., & Van den Abbeele, A. D. (2018). Development and validation of a new MRI simulation technique that can reliably estimate optimal in vivo scanning parameters in a glioblastoma murine model. *PloS one*, *13*(7), e0200611.

Rowley, C. D., Bazin, P. L., Tardif, C. L., Sehmbi, M., Hashim, E., Zaharieva, N., Frey, B. N., & Bock, N. A. (2015). Assessing intracortical myelin in the living human brain using myelinated cortical thickness. Frontiers in neuroscience, 9, 396.

Rowley, C. D., Sehmbi, M., Bazin, P. L., Tardif, C. L., Minuzzi, L., Frey, B. N., & Bock, N. A. (2017). Age-related mapping of intracortical myelin from late adolescence to middle adulthood using T1-weighted MRI. Human brain mapping, 38(7), 3691-3703.

Sacolick, L. I., Wiesinger, F., Hancu, I., & Vogel, M. W. (2010). B1 mapping by Bloch‐Siegert shift. Magnetic resonance in medicine, 63(5), 1315-1322.

Sehmbi, M., Rowley, C. D., Minuzzi, L., Kapczinski, F., Kwiecien, J. M., Bock, N. A., & Frey, B. N. (2019). Age-related deficits in intracortical myelination in young adults with bipolar disorder type I. Journal of psychiatry & neuroscience: JPN, 44(2), 79.

Shams, Z., Norris, D. G., & Marques, J. P. (2019). A comparison of in vivo MRI based cortical myelin mapping using T1w/T2w and R1 mapping at 3T. PloS one, 14(7), e0218089.

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N. E. A., Rapoport J., & Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. Nature, 440(7084), 676.

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE transactions on medical imaging, 17(1), 87-97.

Sled, J. G., & Pike, G. B. (2000). Correction for B1 and B0 variations in quantitative T2 measurements using MRI. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 43(4), 589-593.

Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: a structural description of the human brain. PLoS computational biology, 1(4), e42.

Stanisz, G. J., Odrobina, E. E., Pun, J., Escaravage, M., Graham, S. J., Bronskill, M. J., & Henkelman, R. M. (2005). T1, T2 relaxation and magnetization transfer in tissue at 3T. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 54(3), 507-512.

Stikov, N., Boudreau, M., Levesque, I. R., Tardif, C. L., Barral, J. K., & Pike, G. B. (2015). On the accuracy of T1 mapping: searching for common ground. Magnetic resonance in medicine, 73(2), 514-522.

Stueber, C., Morawski, M., Schäfer, A., Labadie, C., Wähnert, M., Leuze, C., et al. (2014). Myelin and iron concentration in the human brain: a quantitative study of MRI contrast. Neuroimage, 93, 95-106.

Tishler, T. A., Bartzokis, G., Lu, P. H., Raven, E. P., Khanoyan, M., Kirkpatrick, et al. (2018). Abnormal trajectory of intracortical myelination in schizophrenia implicates white matter in disease pathophysiology and the therapeutic mechanism of action of antipsychotics. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3(5), 454-462.

Takao, H., Hayashi, N., & Ohtomo, K. (2014). Effects of study design in multi-scanner voxel-based morphometry studies. Neuroimage, 84, 133-140.

Tsialios, P., Thrippleton, M., Glatz, A., & Pernet, C. (2017). Evaluation of MRI sequences for quantitative T1 brain mapping. In Journal of Physics Conference Series. 931(1).

Van de Moortele, P. F., Auerbach, E. J., Olman, C., Yacoub, E., Uğurbil, K., & Moeller, S. (2009). T1 weighted brain images at 7 Tesla unbiased for Proton Density, T2$*$ contrast and RF coil receive B1 sensitivity with simultaneous vessel visualization. Neuroimage, 46(2), 432-446.

Wang, D., Zuehlsdorff, S., & Larson, A. C. (2009). Rapid 3D radiofrequency field mapping using catalyzed double-angle method. NMR in Biomedicine: An International

Journal Devoted to the Development and Application of Magnetic Resonance In vivo, 22(8), 882-890.

Wang, J., Qiu, M., & Constable, R. T. (2005). In vivo method for correcting transmit/receive nonuniformities with phased array coils. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 53(3), 666-674.

Wang, J., Qiu, M., Yang, Q. X., Smith, M. B., & Constable, R. T. (2005). Measurement and correction of transmitter and receiver induced nonuniformities in vivo. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 53(2), 408-417.

Wang, J., He, L., Zheng, H., & Lu, Z. L. (2014). Optimizing the magnetization-prepared rapid gradient-echo (MP-RAGE) sequence. *PLoS One*, *9*(5), e96899.

Weiskopf, N., Suckling, J., Williams, G., Correia, M. M., Inkster, B., Tait, R., et al. (2013). Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: a multi-center validation. Frontiers in neuroscience, 7, 95.

Yu, M., Linn, K. A., Cook, P. A., Phillips, M. L., McInnis, M., Fava, M., et al. (2018). Statistical harmonization corrects site effects in functional connectivity measurements from multi‐site fMRI data. Human brain mapping, 39(11), 4213-4227.