A CROSS-SECTIONAL, MULTI-SITE STUDY OF T1 WEIGHTED MAGNETIC RESONANCE IMAGE PROPERTIES

by

PHIL JOHN GREER

BS, University of Pittsburgh, 1994

MS, University of Pittsburgh, 1998

Submitted to the Graduate Faculty of

Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Phil John Greer

It was defended on

January, 28, 2013

and approved by

James T Becker, Ph.D., Professor of Psychiatry, Neurology and Psychology,

University of Pittsburgh School of Medicine

Ada O. Youk, Ph.D., Research Assistant Professor, Department of Biostatistics,

Graduate School of Public Health, University of Pittsburgh

Thesis Director: Sati Mazumdar, Ph.D., Professor of Biostatistics and Psychiatry,

Department of Biostatistics,

Graduate School of Public Health, University of Pittsburgh

A CROSS-SECTIONAL, MULTI-SITE STUDY OF T1 WEIGHTED MAGNETIC RESONANCE IMAGE PROPERTIES

Phil John Greer, M.S.

University of Pittsburgh, 2013

Cross sectional Magnetic Resonance Imaging (MRI) studies of the human brain have shown age related changes in morphometric measures as well as signal properties like tissue contrast. Greater access to open data sets and an emphasis on multicenter studies make repeatability of measures important. This study takes a novel dataset and MRI from the IXI open dataset to examine age related changes in MRI signal properties across sites.

We processed T1 weighted MRIs of 1131 subjects from 3 sites using established tissue segmentation tools and region of interest sampling techniques to get regional and global estimates of signal intensity and volume. Signal intensity estimates were used to calculate gray-white contrast values. We created multiple linear regression models of the age related changes in gray-white contrast, within tissue coefficient of variation (CV), and tissue volume while controlling for the effects of gender and site. In addition we compared the contrast results computed using our method to the recent methods provided with FreeSurfer.

Signal intensity decreased with age in two sites while increasing with age in the third. Gray-white contrast showed a decline with age in most regions across all sites. CV increased with age across all regions, sites and tissue types. GM and WM volume decreases with age globally and regionally while CSF volume increased with age. The three methods show strong proportional bias and poor agreement.

It is problematic to make inferences about MRI signal intensity without first normalizing the result. Measures of contrast and CV are both normalized views of signal intensity and both show similar if inverse relationships with age. Volume changes seen here are in agreement with previously published results, but the rates differ by site. Normalizing volume measurements by inter-cranial volume removes the site-specific rate differences. While the gray-white contrast calculation methods show strong correlation, the estimates are not comparable.

Age related MRI signal property changes exist and may affect morphometric measurements. The public health relevance of this research relates to the possibility that MRI contrast may be a biomarker for use in diagnosing neurodegenerative diseases like Alzheimer's that affect elderly populations allowing earlier detection and treatment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This research was supported in part by funds from the National Institute on Aging and the National Heart, Lung, and Blood Institutes. The studies that provided scans to create the template image included: Cardiovascular Health Study – Cognition Study (AG20098), and the Study of Hypertension in the Elderly (NHLBI HL57529). The author also wishes to that the researchers of the IXI - Information eXtraction from Images (EPSRC GR/S21533/02) project for providing open access to this imaging dataset.

The author also wishes to thank the members of his committee, Ada O. Youk, James T. Becker and Sati Mazumdar. Their gentle encouragement helped get this finished in a timely manner. Finally, thanks to faculty members of the University of Pittsburgh who helped in forming arguments presented in this research including; H. Aizenstein, H. M. Gach, J. Price, and J.R. Jennings.

1.0 INTRODUCTION

Magnetic Resonance Imaging (MRI) is a medical imaging technique that allows safe, non-invasive measurements of human tissue *in vivo* with sub-millimeter resolution. As with all medical imaging, the ability to discern different types of tissues and their health status comes down to a matter of signal contrast. The ability to measure and characterize changes in an image property such as signal contrast may help us predict early signs of disease.

1.1 BACKGROUND

MRI works by measuring changes in the activity of atomic nuclei, with the most common element measured being hydrogen (^{1}H) [1]. First, a subject is placed in a static magnetic field (B_0) (the scanner), thereby polarizing the nuclear magnetic spin angular momenta of ^{1}H (spins) in the direction of B_0. Then a radio frequency (RF) pulse is transmitted into the subject, which rocks the spins into a plane that is orthogonal to B_0. Magnetic field gradients are used to spatially encode the spins and acquire an image using a radiofrequency antenna (or coil). The spins relax back to alignment along B_0 using a characteristic spin-lattice relaxation time (T_1). T_1 rises with B_0. The spins dephase in the transverse plane according to the characteristic spin-spin relaxation time (T_2). T_2 is weakly dependent on field strength although it can be affected by magnetic field inhomogeneities that rise with B_0. By changing the sequence of RF and gradient pulses, many

different types of images can be created of the same tissue often showing very different details. This versatility enables MRI to be used not only for medical diagnostics such as finding cancerous tumors or damaged soft tissue, but also in research areas looking at brain anatomy, function, connectivity and in measuring biochemistry using MR spectroscopy.

One such research area is the study of age related changes in the anatomy of the human brain. Previous MRI studies have shown age related changes on global and localized tissue volume[2-11], density[10, 12, 13], cortical thickness[14, 15], tissue microstructure with Diffusion Tensor Imaging[16-22] and more recently MRI signal intensity and contrast [23-26]. Changes in signal intensity and tissue contrast may have fundamental effects on how well tissue segmentation as well as spatial normalization algorithms work. These changes may therefore influence the final volumetric and density measurements. A recent study on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset has shown differences in gray-white contrast due to different scanner hardware, but this study did not look at age related changes in gray-white contrast[27]. If these changes are scanner/sequence specific, they may introduce yet another variable that must be accounted for when combining datasets across multiple acquisition sites. Conversely, changes in tissue contrast may be an early biomarker of some neurodegenerative disease that may be visible before volume loss becomes apparent, and therefore may help in diagnosis[28].

Early methods measuring brain issue signal contrast consisted mainly of region of interest (ROI) analysis using MRI relaxometry[29-32]. MRI relaxometry studies allow careful measurement of imaging parameters such as T1 and T2 signal decay. An early study by Raz et al. showed a decrease in T1 relaxation times in white matter along with reduced differentiation of gray and white matter with increasing age[31]. It is believed that most of this decrease in contrast

is due to changes in white matter due to myelin degradation[29, 31, 33, 34]. While relaxometry is the preferred method for measuring changes in T1 or T2 signal, these studies tend to have long imaging times, relatively poor resolution/coverage and are very labor intensive to analyze. As a result, they often featured comparatively few subjects. More recently, access to large multi-center datasets and automated analysis programs have made it relatively easy to sample signal intensity parameters from large numbers of high-resolution T1 weighted structural MRI[14, 24-27]. The current studies looking at age related contrast changes use FreeSurfer (http://surfer.nmr.mgh.harvard.edu/) to sample Gray Matter (GM) and White Matter (WM) signal intensities across the cerebral cortex creating a cortical map of the ratio of gray matter to white matter. While this is a simple and creative way to measure contrast, it may open itself up to sampling errors due to fluctuations in signal uniformity across the tissue being sampled. In addition, subcortical structures and the cerebellum are mostly ignored.

1.2 OBJECTIVE

In the present study, we wished to explicitly model the age related changes in MRI image properties across multiple sites. By using established tissue segmentation tools[35] and a region of interest approach, we were able to sample T1 Weighted MRI images for volume and signal intensity of Gray Matter, White Matter and Cerebrospinal Fluid (CSF) tissue components. Using the mean and standard deviation of the signal intensities for each tissue type we were able to calculate gray-white contrast as well as the Coefficient of Variation (CV) for each tissue class. We then used multiple linear regression analysis to model the age related changes in signal intensity, gray-white contrast, CV for each tissue class, and tissue volume using gender and site

along with their interaction with age as covariates. Finally, we compared the two current popular sampling methods contained in the FreeSurfer program to the slightly more traditional ROI method used here.

The goals of this study are to discover if there is a decline in gray-white contrast with a method other than the current FreeSurfer based methods, to measure rates of decline in gray-white contrast across scanners and MRI sequences in hopes of assessing repeatability, and to assess the age related changes and repeatability across sites for the CV of each tissue class, Signal Intensity measurements and Volume measurements. Our final goal is to assess the measurement values for gray-white contrast across all methods for agreement.

1.3 PUBLIC HEALTH SIGNIFICANCE

The public health significance of this research relates to the accuracy of MRI as a diagnostic tool. As the life expectancy of our population increases, we become more susceptible to a number of late-life specific diseases such as Alzheimer's disease. Current aging research relies on MRI and a number of processing tools to give us understanding of the changes occurring in the human body as we age. By better understanding basic changes in the MR images themselves, we may be able to use this information to create better processing tools and perhaps aid us in diagnosing these late-life disease states.

2.0 METHODS

2.1 SUBJECTS AND IMAGING

A total of 1,131 images from three sites were used in this study. Subjects ranged in age from 20 to 95 years at scan time with 54.1% of subjects being female. Full summary data for all subjects broken down by scanning site can be found in Table 1. Over 56% of the subjects came from our scanning site at the University of Pittsburgh, while the remainder came from a publicly available image database. The breakdown is as follows:

We were given permission to analyze MRIs from 636 subjects enrolled in one of three different research protocols at the University of Pittsburgh; Cardiovascular Health Study – Cognition Study[36]; and two separate studies of Hypertension in the Elderly[37, 38]. MRIs were obtained for all subjects on a single General Electric (Milwaukee, WI) 1.5 Tesla Signa MR Scanner located at the University of Pittsburgh MR Research Center. Images were acquired using a spoiled gradient-recalled (SPGR) volumetric T1-weighted structural MR scan at time of enrollment [TE=5, TR =25, flip angle = 40 degrees, NEX = 1, slice thickness = 1.5mm/0mm inter-slice] using a standard head coil. All subjects were identified as cognitively normal and healthy using the study-specific criteria.

An additional 495 MRIs were taken from the IXI database (http://www.brain-development.org), which consists of subjects from three scanning locations in the greater London, UK region. For our study we focused on subjects from two sites (Hammersmith Hospital [HH] and Guys Hospital [GH]) due to these cohorts having a more even distribution across the adult age range while the third site had very few elderly subjects. Full details of these scanning sites and relevant scanning parameters can be found in their website. Briefly, both hospitals employed Philips MRI scanners but with differing field strength, (HH = 3T and GH = 1.5T), and used similar T1-FFE scanning sequences [HH: TR/TE/flip angle = 9.6/4.6/8.0, GH: TR/TE/flip angle = 9.8/4.6/8.0] with the only difference being that HH had an Echo Train Length (ETL) of 208 while the ETL at GH was 0.

Table 1. Gender and Age breakdown of all subjects by scanning site

	Obs	Mean	Std. Dev.	Min	Max
University of Pittsburgh					
Female	343	74.04	9.937	36	92
Male	293	71.03	11.781	35	95
Guy's Hospital					
Female	175	51.88	15.148	21	80
Male	139	49.47	16.823	20	86
Hammersmith Hospital					
Female	94	50.06	17.055	21	82
Male	87	44.54	15.972	20	79
TOTAL	**1131**	**62.55**	**17.731**	**20**	**95**

2.2 IMAGE PROCESSING

All images were processed using FSL 4.1[35]. First, all images were labeled with correct orientation information and placed in axial orientation in radiologic format. Images were skull stripped using the BET (Brain Extraction Tool) program in FSL then bias corrected as well as segmented into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) tissue class images using FAST (FMRIB's Automated Segmentation Tool). A threshold of 50% was then applied to the tissue class images to remove edge voxels that have a lower probability of belonging to the tissue type in the image. This thresholding turns the images into binary masks where the tissue class is encoded with a value of 1 and all other areas are encoded as 0. Examples of segmentation results can be found in Figure 1.

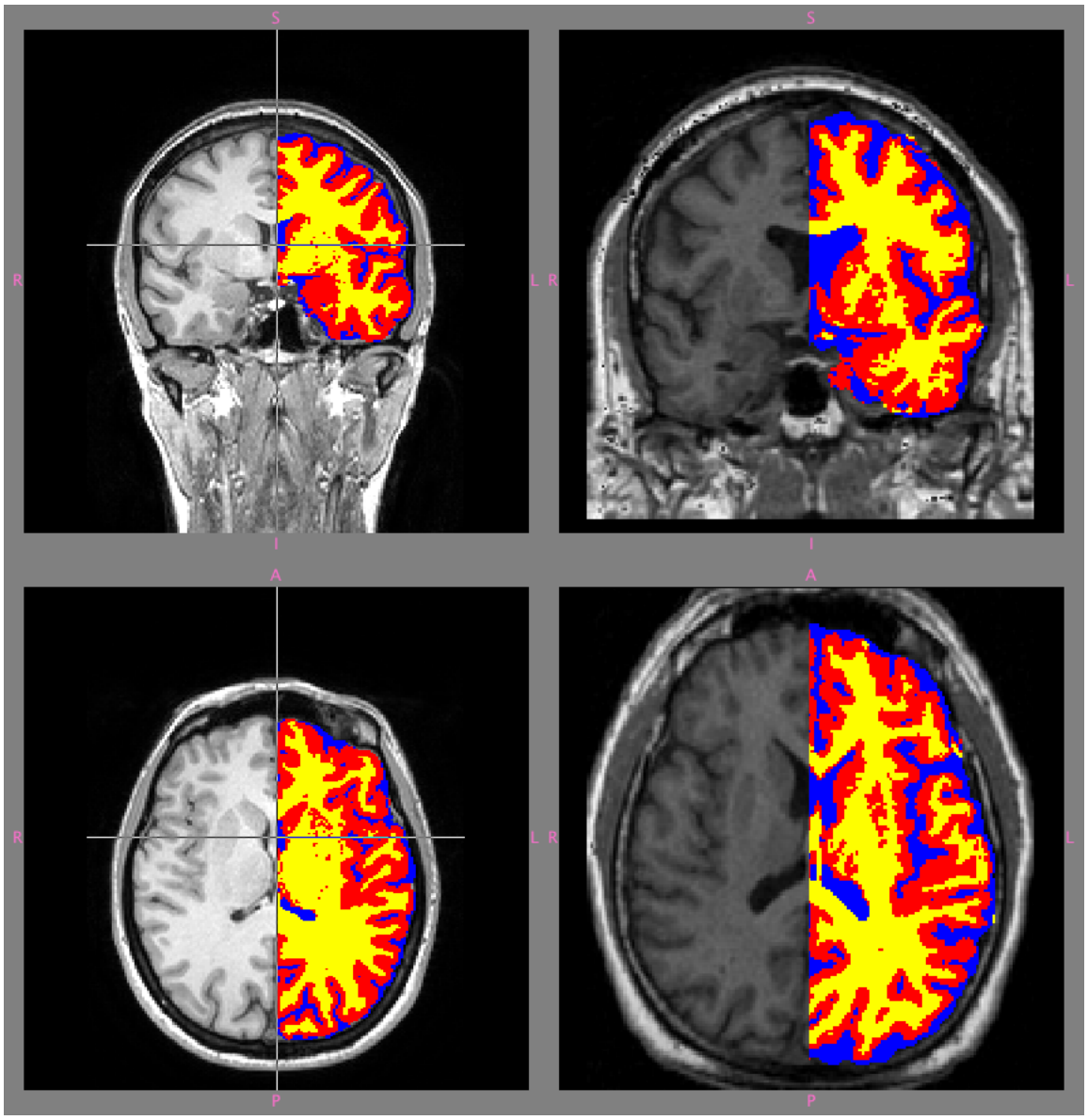

Figure 1. Image segments over-layed on T1 MRI in subject space.

Left 35yo female from Guys Hospital, Right 63yo female from the University of Pittsburgh. Segmented images are color coded to tissue type: Blue= CSF, Red=Gray Matter, and Yellow=White Matter.

Skull stripped brains were then aligned to the MNI152 standard space provided by the Montreal Neurologic Institute (MNI) using the affine registration tool FLIRT (FMRIB's Linear Image Registration Tool). The standard ROIs used for this study were taken from the MNI

Structural atlas with a probability threshold of zero to maximize the amount of each separate tissue class found within the ROI. A list of these regions and labels can be found in Figure 2. These ROIs were then mapped back from standard space to subject space using the inverse of the FLIRT transform calculated previously, then multiplied by each subject space, tissue component image using the FSL image manipulation program fslmaths, thus creating three separate ROI files for each region (GM, WM and CSF). Finally, these regional tissue ROI images were made binary using the same 50% threshold algorithm as before to clean up any partial volume averaging created by the inverse FLIRT transform from standard to subject space. Individual bias-corrected MRI scans were converted to 8-bit (dynamic range of 0-255) and then sampled using the binarized, regional tissue ROIs to get mean signal intensities and standard deviation (SD) for each tissue class per region using the FSL sampling program fslstats. In addition global and regional volumes for GM, WM and CSF were calculated using fslstats. Inter-cranial volume (ICV) was calculated using the sum of the whole brain GM, WM and CSF volumes. Several deep brain structures (Caudate, Putamen, and Thalamus) contained little to no CSF signal and were dropped from further analysis requiring CSF measures.

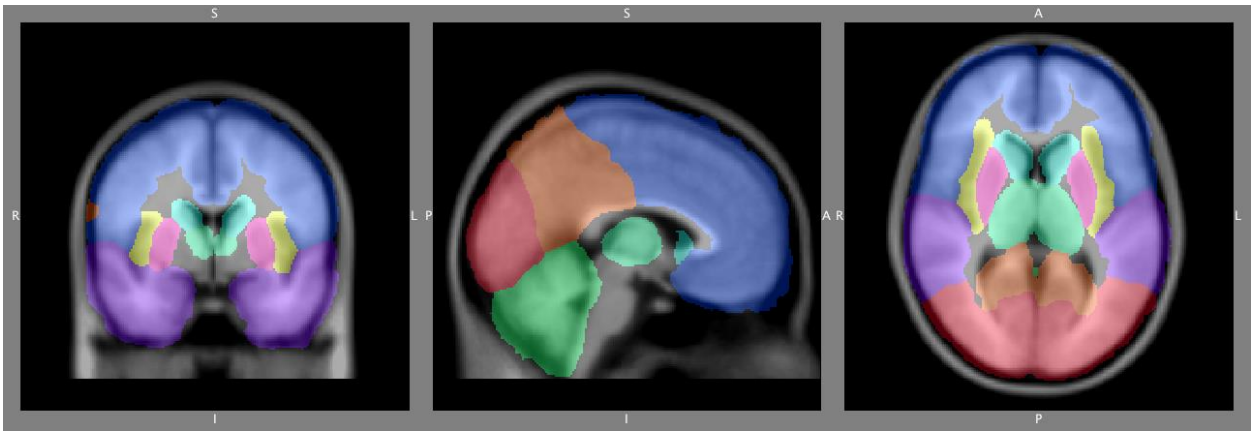

Figure 2. MNI Regions set at probability threshold 0 over-layed on MNI atlas image.

Regions seen here: Frontal=blue, Cerebellum=green, Thalamus=light green, Caudate=aqua, Putamen=pink, Insula=yellow, Parietal=orange, Temporal=purple, Occipital=red

2.3 CONTRAST CALCULATION

Between tissue contrast was calculated using a variation of the Michelson formula [39]

$$\left(\frac{\bar{X}_{tissue1} - \bar{X}_{tissue2}}{\bar{X}_{tissue1} + \bar{X}_{tissue2}} \right) * 100$$

This formula provides a range of values from -100, where the value for tissue1 is zero, to +100, where the value for tissue2 is zero. In the case of gray-white contrast, tissue1 is WM and tissue2 is GM. While this value is highly correlated with the white/gray ratio as calculated in prior studies [24, 26, 27](R=0.998), we feel that this value, which represents a % difference, may be easier to understand than a simple ratio with the added benefit that there will be less likelihood of finding a zero in the denominator. Contrast was calculated for gray-white both globally and regionally for each ROI.

2.4 STATISTICAL ANALYSIS

Extracted ROI signal intensity data were analyzed using STATA v12 (College Station, TX). Coefficient of Variation (CV) for each tissue type was calculated using the SD of the signal intensity for the tissue type divided by the mean signal intensity for that tissue type. CVs were calculated for all tissue classes per region.

Multiple linear regression analysis was performed to assess the relationships between global and regional measures of: signal intensity, gray-white contrast, CV for each tissue class, and tissue volumes (with and without normalization to inter-cranial volume) with three main predictor variables. These variables were age (centered to the overall mean), gender, and

acquisition site along with all possible second and third order interactions. A model key with an explanation of all variables can be found in Appendix A.

2.5 TEST OF AGREEMENT

A sample of 106 subjects from the initial group was selected at random for comparing methods of contrast sampling. In addition to the segmentation processing described earlier, these 106 subjects were processed using the FreeSurfer image analysis suite, which is documented and freely available for download online (http://surfer.nmr.mgh.harvard.edu/). Measurements of cortical thickness were obtained by reconstructing representations of the WM/GM boundary and the cortical surface and then calculating the distance between these surfaces at each vertex across the cortical mantle. The technical details of these procedures are described in prior publications[40-48]. Importantly, the thickness maps are created using three-dimensional spatial intensity gradients across tissue classes and are therefore not simply reliant on absolute signal intensity. The maps are not restricted to the voxel resolution of the original data and are capable of detecting sub-millimeter differences in cortical thickness between groups[41]. Signal intensities were sampled using the methods of both Salat et al 2009[24] and Westlye et al 2009[26] using the ptcsurfcon program distributed with FreeSurfer v5.1. Each method samples the signal intensity at a single point in both WM and GM at a set distance from each vertex of the GM/WM boundary. The method used by Salat samples WM signal intensity 1mm into the WM while sampling GM signal intensity at a point 35% into the thickness of the GM cortical ribbon. The method of Westlye uses a fixed distance of 0.2mm from GM/WM boundary for measuring both GM and WM signal intensity. Unlike the papers published by Salat and Westlye, which

report simplified GM/WM or WM/GM ratios respectively, the ptcsurfcon program calculates a contrast measure similar to the one reported here and creates mean values based on atlas regions of the cortex. The only difference between the measure of contrast calculated by the pctsurfcon program and the one used in this paper is that the formula used by FreeSurfer v5.1 contains a multiplicative factor of two. For this reason all corresponding segmentation based contrast measures were multiplied by two to more closely match the FreeSurfer results. The default atlas regions in FreeSurfer were combined into “lobular” regions to more closely match the regions from the MNI atlas found in FSL. This left five cortical regions for comparison: Frontal, Insula, Occipital, Parietal and Temporal.

Comparisons of agreement were performed using the method of Bland and Altman[49], with additional measures of correlation, and regression analysis to calculate slopes. All tests were done using STATA v12.

3.0 RESULTS

3.1 SIGNAL INTENSITIES

As seen in Figure 3, the first two sites (Pittsburgh and Guys Hospital) show global signal intensities for all tissue components decrease significantly with age, while at the third site (Hammersmith Hospital), global signal intensities for all tissue components increase with age. Regional analysis was not done for signal intensities.

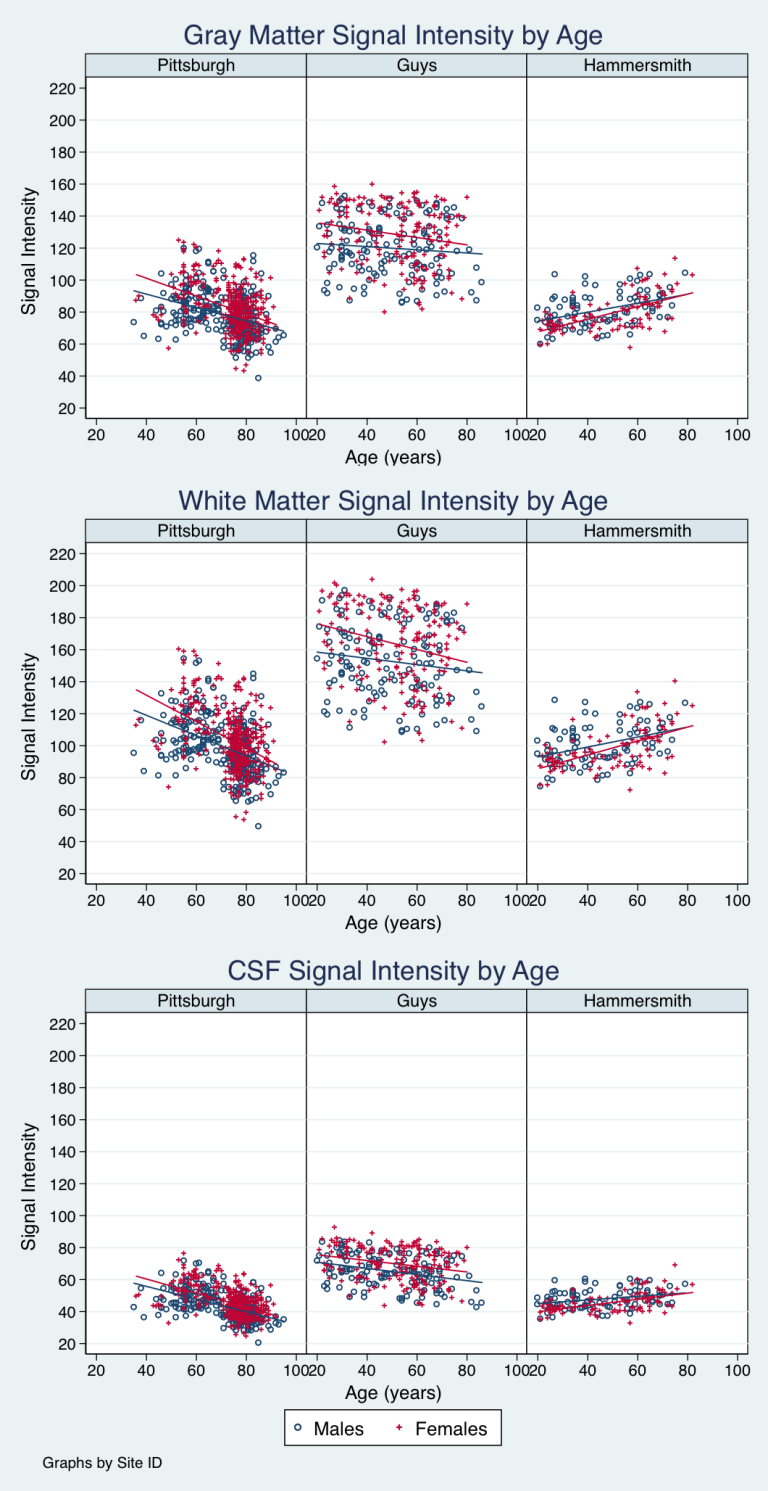

Figure 3. Signal intensities for each tissue class by age.

All graphs are separated by gender and site.

3.2 CONTRAST

Models for whole brain and other regions can be found in Table 2. Gray-white tissue contrast showed significant relationships with age, gender, and site, along with significant interactions between age and gender as well as age and site. No 3-way interactions were found. Visual representation of this relationship can be found in Figure 4. Whole brain gray-white contrast significantly decreases with age, while the specific rate of decrease depends on both gender and site. Males showed a somewhat higher amount of gray-white contrast than females at the mean age of the sample and showed a slower decline with age. This relationship held true for most regional samples as well, the exceptions being that gray-white contrast showed no significant age by gender interaction in caudate and no age by site interaction in Cerebellum as seen in Table 2. In addition, the putamen showed a positive relationship with each year of age.

Table 2. Multiple regression results of age, gender and site on Gray-White Contrast.

Models shown for whole brain and 9 sub-regions. See model key for full explanation of variable names.

	Whole Brain	**Caudate**	**Cerebellum**	**Frontal**	**Insula**	**Occipital**	**Parietal**	**Putamen**	**Temporal**	**Thalamus**
	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)
1.Gender	0.1143	0.1705	0.1350	0.1487	0.0982**	0.1923	0.0890*	0.1496	0.0990**	0.1178**
	(0.0316)	(0.0452)	(0.0341)	(0.0340)	(0.0380)	(0.0325)	(0.0346)	(0.0310)	(0.0310)	(0.0369)
AgeC	-0.0579	-0.0752	-0.0129	-0.0665	-0.0800	-0.0418	-0.0680	0.0121	-0.0577	-0.0394
	(0.0021)	(0.0030)	(0.0023)	(0.0023)	(0.0025)	(0.0022)	(0.0023)	(0.0021)	(0.0021)	(0.0024)
1.Gender#AgeC	0.0106	0.0041†	0.0104	0.0105	0.0095	0.0082	0.0116	0.0063	0.0090	0.0087
	(0.0018)	(0.0025)	(0.0019)	(0.0019)	(0.0021)	(0.0018)	(0.0019)	(0.0017)	(0.0017)	(0.0021)
1.Site	-0.5082	-1.7160	-0.5462	-0.5269	-0.2417	-0.5093	-0.7742	-0.5076	-0.3023	-0.1042†
	(0.0466)	(0.0668)	(0.0504)	(0.0502)	(0.0561)	(0.0481)	(0.0511)	(0.0458)	(0.0458)	(0.0545)
2.Site	-1.7285	-3.0804	-1.5117	-1.8213	-2.3012	-1.7210	-1.8854	-1.5729	-2.0085	-0.3854
	(0.0599)	(0.0859)	(0.0647)	(0.0646)	(0.0721)	(0.0617)	(0.0657)	(0.0588)	(0.0588)	(0.0700)
1.Site#AgeC	0.0250	0.0529	-0.0048†	0.0296	0.0448	0.0240	0.0337	-0.0198	0.0265	0.0103
	(0.0027)	(0.0038)	(0.0029)	(0.0029)	(0.0032)	(0.0027)	(0.0029)	(0.0026)	(0.0026)	(0.0031)
2.Site#AgeC	0.0322	0.0678	-0.0009†	0.0376	0.0624	0.0285	0.0384	-0.0220	0.0380	0.0188
	(0.0030)	(0.0043)	(0.0033)	(0.0033)	(0.0036)	(0.0031)	(0.0033)	(0.0030)	(0.0030)	(0.0035)
constant	12.1041	10.2223	10.2313	12.4887	12.1138	12.3572	12.4905	7.8146	12.8321	9.3627
	(0.0329)	(0.0471)	(0.0355)	(0.0354)	(0.0396)	(0.0339)	(0.0360)	(0.0323)	(0.0323)	(0.0384)
R-sqr	0.614	0.652	0.431	0.633	0.711	0.533	0.599	0.581	0.693	0.357
dfres	1123	1124	1125	1123	1123	1123	1123	1123	1123	1123

* $p<0.05$, ** $p<0.01$, † not significant

All betas (β) significant at $p<0.001$ unless otherwise marked

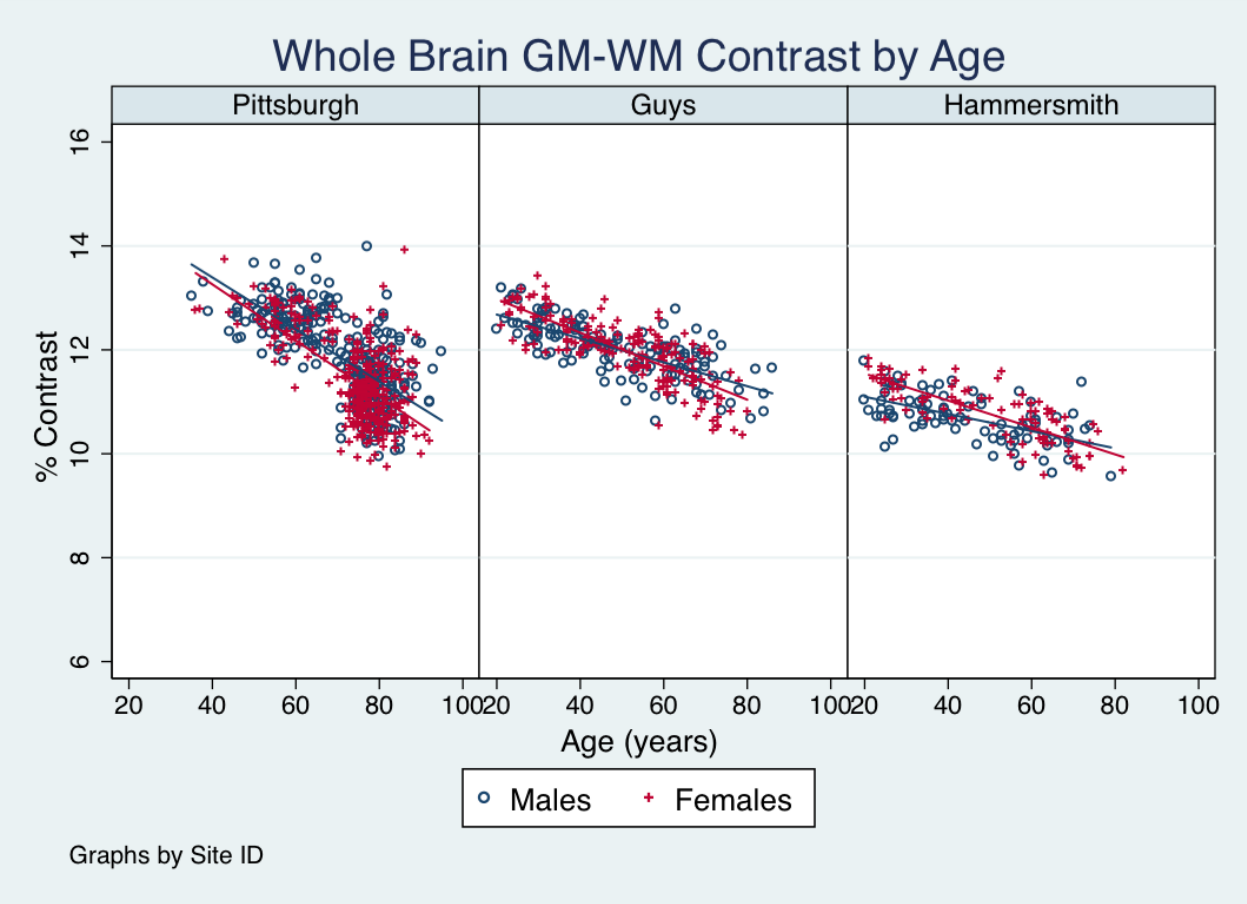

Figure 4. Whole brain Gray White Contrast by age, separated by gender and scanning site.

3.3 COEFFICIENT OF VARIATION

Models for whole brain and other regions can be found in Tables 3, 4, and 5. The CV of all tissue classes showed significant relationships with age, gender, and site, as well as significant interactions between age and gender and age and site. No 3-way interactions were found. Visual representation of this relationship for GM and WM can be found in Figure 5. The CVs for each tissue class increased with age, but exact rates were dependent on gender and site. Males showed

a somewhat higher CV in WM and GM than females at the mean age of the sample and showed a greater increase with each year of age. In CSF CV, there were borderline to no significant difference between males and females.

Of note, due to a lack of CSF voxels in subcortical ROIs, Caudate, Putamen, and Thalamus regions were not analyzed for CSF CV. For the most part regional models were consistent with whole brain models.

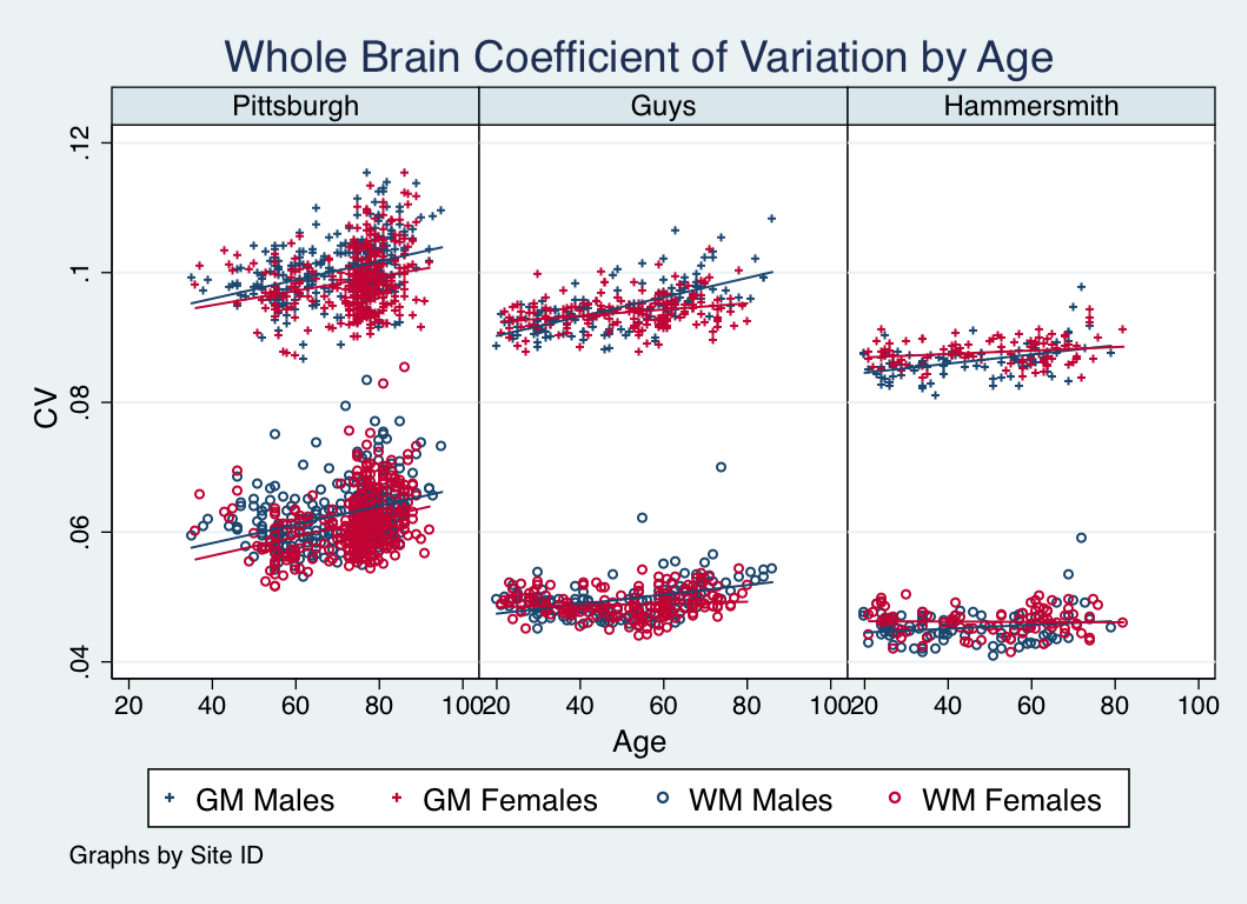

Figure 5. Whole brain gray matter (pluses) and white matter (open circles) coefficient of variation by age, separated by gender (blue=male, maroon=female) and scanning site.

Table 3. Multiple regression results of age, gender and site on Gray Matter Coefficient of Variation.

Models shown for whole brain and 9 sub-regions. See model key for full explanation of variable names.

	Whole Brain	**Caudate**	**Cerebellum**	**Frontal**	**Insula**	**Occipital**	**Parietal**	**Putamen**	**Temporal**	**Thalamus**
	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)
1.Gender	0.0013	0.0011**	0.0016	0.0012	0.0016	0.0014	0.0012	0.0014**	0.0013	0.0015
	(0.0002)	(0.0004)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0004)	(0.0002)	(0.0004)
AgeC	0.0001	0.0004	0.0003	0.0000†	-0.0001	0.0002	0.0001	0.0003	0.0002	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
1.Gender#AgeC	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000†	0.0001	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
1.Site	-0.0030	-0.0073	-0.0001†	-0.0043	-0.0066	0.0003†	-0.0046	-0.0020**	-0.0045	0.0068
	(0.0003)	(0.0006)	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0006)	(0.0003)	(0.0006)
2.Site	-0.0104	-0.0070	-0.0068	-0.0121	-0.0153	-0.0083	-0.0121	-0.0078	-0.0106	0.0057
	(0.0004)	(0.0008)	(0.0006)	(0.0004)	(0.0004)	(0.0005)	(0.0004)	(0.0008)	(0.0004)	(0.0008)
1.Site#AgeC	-0.0000†	-0.0002	-0.0002	0.0000*	0.0001	-0.0001	0.0000†	-0.0003	-0.0001	-0.0001*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
2.Site#AgeC	-0.0001**	-0.0002	-0.0002	0.0000†	0.0001	-0.0002	-0.0001*	-0.0003	-0.0002	-0.0001†
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
constant	0.0978	0.0805	0.0930	0.0989	0.0922	0.0997	0.0980	0.0565	0.0956	0.0887
	(0.0002)	(0.0004)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0004)	(0.0002)	(0.0004)
R-sqr	0.623	0.578	0.504	0.655	0.727	0.621	0.669	0.373	0.697	0.142
dfres	1123	1123	1123	1123	1123	1123	1123	1123	1123	1123

* $p<0.05$, ** $p<0.01$, † not significant

All betas (β) significant at $p<0.001$ unless otherwise marked

Table 4. Multiple regression results of age, gender and site on White Matter Coefficient of Variation.

Models shown for whole brain and 9 sub-regions. See model key for full explanation of variable names.

	Whole Brain	**Caudate**	**Cerebellum**	**Frontal**	**Insula**	**Occipital**	**Parietal**	**Putamen**	**Temporal**	**Thalamus**
	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)
1.Gender	0.0011	0.0018	0.0009	0.0013	0.0010	0.0008	0.0010	0.0015	0.0010	0.0014
	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0002)
AgeC	0.0001	0.0000*	-0.0001	0.0001	0.0000†	0.0001	0.0002	0.0001	0.0001	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
1.Gender#AgeC	0.0000	0.0001	0.0000**	0.0001	0.0001	0.0000	0.0000**	0.0000**	0.0001	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
1.Site	-0.0109	-0.0022	-0.0084	-0.0128	-0.0084	-0.0115	-0.0106	-0.0065	-0.0142	-0.0056
	(0.0003)	(0.0005)	(0.0003)	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0003)
2.Site	-0.0145	-0.0062	-0.0132	-0.0159	-0.0128	-0.0149	-0.0144	-0.0114	-0.0164	-0.0130
	(0.0004)	(0.0006)	(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0005)	(0.0004)	(0.0005)	(0.0004)
1.Site#AgeC	-0.0001	-0.0001**	0.0001	-0.0001	-0.0001	-0.0001	-0.0002	-0.0002	-0.0001**	-0.0000†
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
2.Site#AgeC	-0.0001	-0.0001*	0.0001	-0.0001	-0.0001	-0.0002	-0.0002	-0.0001	-0.0001	-0.0000†
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
constant	0.0601	0.0592	0.0648	0.0592	0.0511	0.0606	0.0576	0.0570	0.0656	0.0578
	(0.0002)	(0.0004)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0002)
R-sqr	0.795	0.210	0.688	0.777	0.708	0.758	0.772	0.706	0.797	0.774
dfres	1123	1123	1123	1123	1123	1123	1123	1123	1123	1123

* $p<0.05$, ** $p<0.01$, † not significant

All betas (β) significant at $p<0.001$ unless otherwise marked

Table 5. Multiple regression results of age, gender and site on cerebro-spinal fluid Coefficient of Variation.

Models shown for whole brain and 6 sub-regions. See model key for full explanation of variable names.

	Whole Brain	**Cerebellum**	**Frontal**	**Insula**	**Occipital**	**Parietal**	**Temporal**
	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)
1.Gender	-0.0052**	-0.0162	-0.0028†	0.0005†	-0.0031†	-0.0022†	-0.0135
	(0.0020)	(0.0027)	(0.0017)	(0.0018)	(0.0032)	(0.0025)	(0.0027)
AgeC	0.0009	-0.0000†	0.0011	0.0021	-0.0001†	0.0013	-0.0005**
	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
1.Gender#AgeC	0.0002*	0.0003†	0.0002*	0.0003**	0.0007	0.0002†	0.0003*
	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0002)
1.Site	-0.0058*	-0.0021†	0.0017†	0.0491	-0.0053†	0.0171	-0.0418
	(0.0029)	(0.0040)	(0.0025)	(0.0026)	(0.0047)	(0.0037)	(0.0040)
2.Site	-0.0153	-0.0272	-0.0029†	0.0014†	-0.0028†	-0.0219	0.0178
	(0.0038)	(0.0051)	(0.0032)	(0.0034)	(0.0060)	(0.0048)	(0.0051)
1.Site#AgeC	-0.0012	-0.0008	-0.0010	-0.0020	-0.0008**	-0.0014	-0.0008
	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0003)	(0.0002)	(0.0002)
2.Site#AgeC	-0.0023	-0.0019	-0.0018	-0.0041	-0.0015	-0.0017	-0.0023
	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0003)
constant	0.3315	0.3582	0.2936	0.2267	0.3026	0.2783	0.3860
	(0.0021)	(0.0028)	(0.0018)	(0.0019)	(0.0033)	(0.0026)	(0.0028)
R-sqr	0.148	0.108	0.172	0.411	0.056	0.142	0.373
dfres	1123	1123	1123	1123	1123	1123	1123

* $p<0.05$, ** $p<0.01$, † not significant

All betas (β) significant at $p<0.001$ unless otherwise marked

3.4 VOLUMES

Models for various regions can be found in Tables 6, 7, and 8. Volume measurements for all three classes showed significant relationships with age, gender, and site, along with significant interactions between age and site. GM and WM volumes also showed weak, but significant relationship with the age by gender interaction term, but CSF did not. No 3-way interactions were found. Overall GM and WM volume significantly decreases with age, while the specific rate of decrease depends on both gender and site as seen in Figure 6a. CSF volumes significantly increased with age, but showed no age by gender interaction. Males had significantly increased volumes for all tissue types compared to the females in the study, but showed only slightly greater rates of volume decline with age in the case of GM and WM, and no difference in the rate of volume increases in the case of CSF.

Normalizing the tissue compartment volumes by the inter-cranial volume (ICV) rendered both interaction terms (site by age and gender by age) insignificant in each tissue compartment model leaving only main effects as seen in Figure 6b. Estimates from the ICV normalized models can be found in Table 9. Of particular note, normalized WM volume did not fit the model well as can be seen with by its R^2 value of 0.052.

Table 6. Multiple regression results of age, gender and site on Gray Matter Volume.

Models shown for whole brain and 9 sub-regions. See model key for full explanation of variable names.

	Whole Brain	**Caudate**	**Cerebellum**	**Frontal**	**Insula**	**Occipital**	**Parietal**	**Putamen**	**Temporal**	**Thalamus**
	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)
1.Gender	44439.57	353.45	4418.04	12848.48	1137.91	5960.92	8011.36	580.41	10634.72	540.10
	(2723.85)	(71.90)	(533.90)	(906.47)	(97.45)	(383.58)	(594.98)	(73.98)	(602.33)	(54.83)
AgeC	-3591.70	-73.89	-389.01	-1288.52	-85.66	-437.99	-778.08	-29.69	-751.81	-35.54
	(180.95)	(4.78)	(35.47)	(60.22)	(6.47)	(25.48)	(39.53)	(4.91)	(40.01)	(3.64)
1.Gender#AgeC	-359.15*	-18.11	-59.76*	-110.12*	-18.96	-26.05†	-47.06†	-8.33*	-82.01*	-1.66†
	(152.83)	(4.03)	(29.96)	(50.86)	(5.47)	(21.52)	(33.38)	(4.15)	(33.80)	(3.08)
1.Site	-1768.03†	-359.81	9058.98	45.79†	-1133.37	-2690.06	-2857.23**	-1485.74	3202.39	-3177.05
	(4022.78)	(106.18)	(788.50)	(1338.74)	(143.92)	(566.49)	(878.70)	(109.26)	(889.56)	(80.97)
2.Site	-21521.86	-1438.34	-163.86†	4703.28**	-1857.42	-6171.23	-3392.30**	-3748.45	-3224.25**	-4138.70
	(5169.35)	(136.44)	(1013.23)	(1720.30)	(184.94)	(727.95)	(1129.15)	(140.40)	(1143.10)	(104.05)
1.Site#AgeC	1109.06	50.79	-46.49†	392.96	38.95	153.92	208.08	51.54	354.26	56.50
	(230.16)	(6.08)	(45.11)	(76.60)	(8.23)	(32.41)	(50.27)	(6.25)	(50.90)	(4.63)
2.Site#AgeC	881.81	61.94	81.78†	271.52**	0.45†	166.88	80.80†	68.12	264.74	50.35
	(260.29)	(6.87)	(51.02)	(86.62)	(9.31)	(36.65)	(56.86)	(7.07)	(57.56)	(5.24)
constant	512786.56	6122.52	56460.21	155691.53	13927.68	65564.27	103133.14	6947.14	103215.04	8351.45
	(2835.26)	(74.84)	(555.73)	(943.54)	(101.44)	(399.26)	(619.31)	(77.01)	(626.96)	(57.07)
R-sqr	0.638	0.363	0.575	0.693	0.409	0.505	0.626	0.595	0.622	0.791
dfres	1123	1123	1123	1123	1123	1123	1123	1123	1123	1123

* $p<0.05$, ** $p<0.01$, † not significant

All betas (β) significant at $p<0.001$ unless otherwise marked

Table 7. Multiple regression results of age, gender and site on White Matter Volume.

Models shown for whole brain and 9 sub-regions. See model key for full explanation of variable names.

	Whole Brain	**Caudate**	**Cerebellum**	**Frontal**	**Insula**	**Occipital**	**Parietal**	**Putamen**	**Temporal**	**Thalamus**
	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)
1.Gender	69163.58	50.87†	8720.96	18989.56	1198.34	8246.00	14104.28	2121.94	9825.63	1577.74
	(2984.20)	(75.85)	(537.83)	(958.40)	(128.49)	(432.93)	(647.62)	(114.88)	(436.46)	(126.52)
AgeC	-1776.90	-80.21	-126.39	-110.92†	-10.66†	-71.56*	-367.97	-22.95**	-28.98†	-172.61
	(198.25)	(5.04)	(35.73)	(63.67)	(8.54)	(28.76)	(43.02)	(7.63)	(29.00)	(8.41)
1.Gender#AgeC	-451.34**	-10.40*	-45.55†	-122.10*	3.86†	-57.51*	-66.31†	-4.20†	-17.48†	-32.68
	(167.44)	(4.26)	(30.18)	(53.78)	(7.21)	(24.29)	(36.34)	(6.45)	(24.49)	(7.10)
1.Site	-10778.20*	458.87	-12553.92	-1615.06†	969.96	195.89†	961.04†	1108.66	-4165.88	2794.57
	(4407.28)	(112.02)	(794.30)	(1415.44)	(189.76)	(639.39)	(956.44)	(169.67)	(644.60)	(186.86)
2.Site	-10317.89†	1539.07	-3864.44	-13607.17	1989.23	492.52†	-2269.01†	2754.80	-1176.04†	3923.47
	(5663.45)	(143.94)	(1020.69)	(1818.87)	(243.85)	(821.63)	(1229.05)	(218.03)	(828.32)	(240.12)
1.Site#AgeC	1149.63	22.97	262.33	72.47†	1.65†	43.74†	224.98	-21.83*	24.83†	49.83
	(252.16)	(6.41)	(45.45)	(80.98)	(10.86)	(36.58)	(54.72)	(9.71)	(36.88)	(10.69)
2.Site#AgeC	606.96*	1.72†	40.92†	-38.55†	17.31†	-53.00†	204.76	-48.07	-32.10†	48.16
	(285.17)	(7.25)	(51.39)	(91.58)	(12.28)	(41.37)	(61.89)	(10.98)	(41.71)	(12.09)
constant	529721.78	3507.90	65369.36	138475.34	9261.07	58642.53	100133.64	15239.41	64757.65	15406.17
	(3106.26)	(78.95)	(559.82)	(997.60)	(133.75)	(450.64)	(674.10)	(119.58)	(454.31)	(131.70)
R-sqr	0.443	0.661	0.407	0.305	0.174	0.306	0.413	0.527	0.35	0.802
dfres	1123	1123	1123	1123	1123	1123	1123	1123	1123	1123

* $p<0.05$, ** $p<0.01$, † not significant

All betas (β) significant at $p<0.001$ unless otherwise marked

Table 8. Multiple regression results of age, gender and site on cerebrospinal fluid Volume.

Models shown for whole brain and 6 sub-regions. See model key for full explanation of variable names.

	Whole Brain	**Cerebellum**	**Frontal**	**Insula**	**Occipital**	**Parietal**	**Temporal**
	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)	β/(se)
1.Gender	39967.71	2944.08	12011.22	1034.87	3599.14	8804.84	5761.44
	(2042.29)	(199.31)	(680.17)	(78.78)	(223.31)	(481.39)	(341.68)
AgeC	470.24	-32.89*	-293.50	35.81	-27.42†	169.54	23.84†
	(135.67)	(13.24)	(45.18)	(5.23)	(14.83)	(31.98)	(22.70)
1.Gender#AgeC	164.73†	-7.86†	-2.03†	0.83†	20.82†	24.15†	29.39†
	(114.59)	(11.18)	(38.16)	(4.42)	(12.53)	(27.01)	(19.17)
1.Site	-23366.31	-1541.13	-11762.12	-284.11*	-1258.26	-3626.45	-3622.13
	(3016.20)	(294.35)	(1004.52)	(116.35)	(329.79)	(710.95)	(504.62)
2.Site	-14897.52	-404.67†	-5386.95	-1376.00	-1554.64	-2064.45*	-2440.42
	(3875.88)	(378.25)	(1290.83)	(149.51)	(423.79)	(913.59)	(648.45)
1.Site#AgeC	652.47	70.49	584.71	-6.23†	122.24	127.51**	90.73**
	(172.57)	(16.84)	(57.47)	(6.66)	(18.87)	(40.68)	(28.87)
2.Site#AgeC	977.28	119.32	729.30	3.72†	148.19	177.73	163.20
	(195.16)	(19.05)	(65.00)	(7.53)	(21.34)	(46.00)	(32.65)
constant	259951.65	24971.36	86689.24	4268.81	26177.56	54944.73	41861.29
	(2125.82)	(207.46)	(707.99)	(82.00)	(232.44)	(501.08)	(355.66)
R^2	0.473	0.224	0.407	0.404	0.32	0.446	0.364
dfres	1123	1123	1123	1123	1123	1123	1123

* $p<0.05$, ** $p<0.01$, † not significant

All betas (β) significant at $p<0.001$ unless otherwise marked

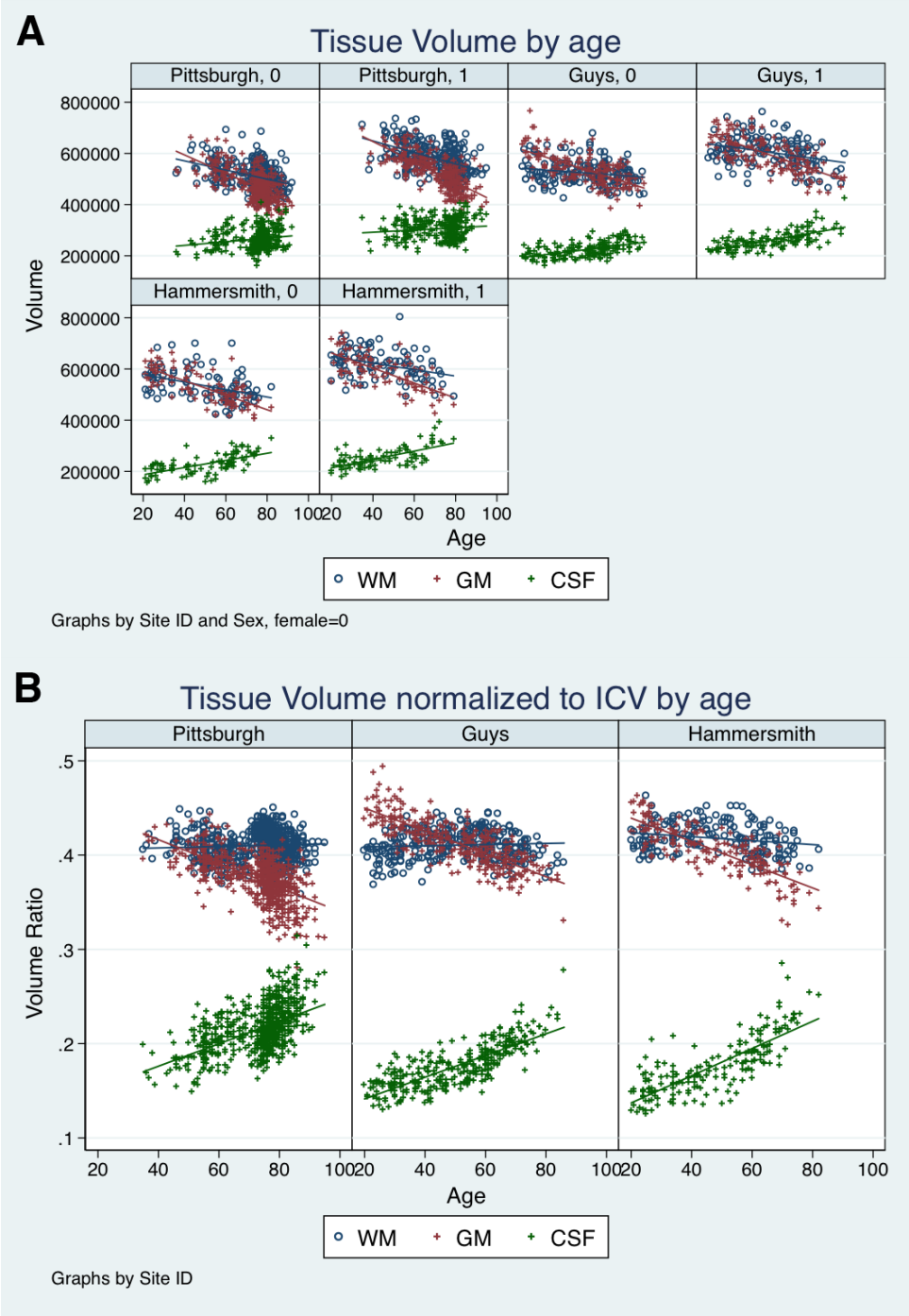

Figure 6. Tissue volume changes by age.

A) Separate graphs for gender as well as scanning site. B) WM, GM and CSF normalized to Inter-cranial Volume (ICV) and separated by scanning site. Due to very small difference between genders, male and female combined for this graph.

Table 9. Whole Brain Volume models normalized to inter-cranial volume

	Gray Matter	White Matter	CSF
	β/(se)	β/(se)	β/(se)
1.Gender	-0.0117	0.0043	0.0074
	(0.0010)	(0.0009)	(0.0011)
AgeC	-0.0013	0.0000†	0.0013
	(0.0000)	(0.0000)	(0.0000)
1.Site	0.0093	0.0009†	-0.0102
	(0.0013)	(0.0013)	(0.0016)
2.Site	-0.0018†	0.0082	-0.0064
	(0.0016)	(0.0015)	(0.0019)
constant	0.3933	0.4080	0.1987
	(0.0009)	(0.0008)	(0.0010)
R^2	0.711	0.052	0.649
dfres	1126	1126	1126

* p<0.05, ** p<0.01, † not significant

All betas (β) significant at p<0.001 unless otherwise marked

3.5 AGREEMENT

Four subjects failed to process using FreeSurfer and were excluded from the analysis leaving a total of 102 subjects. Table 10 shows summary statistics for the contrast measures of each region by sampling method. Figure 7 shows a scatter plot of the pairwise combinations of results. As the information in Table 10 and Figure 7 shows, contrast measures for the Westlye method are 3.5x to 4x less than the contrast calculate by the Salat method while the contrast measures of the Salat method are roughly 10-15% lower than the Segmented ROI method.

Table 11 shows correlation coefficients with 95% confidence intervals for all methods and all regions. In addition, regression equations are shown. While all pairwise combinations show significant correlations, the methods of Salat and Westlye show the highest pairwise correlation.

Figure 8 shows Bland-Altman plots for a select number of regions comparing the Westlye method to both the Segmented ROI method and the Salat method. Both comparisons show strong proportional bias. Figure 9 shows that scaling the Westlye method by a factor of 4 brings its' contrast measures more inline with the other two methods.

Figures 10, 11 and 12 show Bland-Altman plots for all regions and for all pairwise combinations. Table 12 summarizes the bias, standard deviation and percent error. In Table 13, results of tests of skewness and kurtosis show that all pairwise differences are not distributed normally. The results in Table 13 taken with information from Figures 7 and 9 show that the differences are greater with increased contrast. Even with the scaled Westlye contrast results, a large amount of bias and a high percent error remain.

Table 10. Summary statistics for each region by method.

Mean, SD, MIN and MAX are across all observations for selected region and method.

	Obs	Mean Contrast	Standard Deviation	Min Contrast	Max Contrast
SegROI					
Frontal	102	24.267	2.045	20.093	28.897
Insula	102	23.087	2.602	18.102	29.466
Occipital	102	23.838	1.785	20.159	27.785
Parietal	102	23.962	1.941	19.551	28.845
Temporal	102	24.794	2.077	20.489	29.216
Westlye					
Frontal	102	5.586	0.576	4.472	8.047
Insula	102	4.875	0.506	3.888	6.101
Occipital	102	5.592	0.834	4.458	10.466
Parietal	102	5.725	0.599	4.548	8.957
Temporal	102	5.956	0.633	4.578	7.638
Salat					
Frontal	102	20.728	2.056	16.229	25.273
Insula	102	19.226	1.960	15.646	23.821
Occipital	102	19.125	2.231	15.636	30.410
Parietal	102	20.650	2.034	16.027	26.780
Temporal	102	22.448	2.372	18.155	28.330

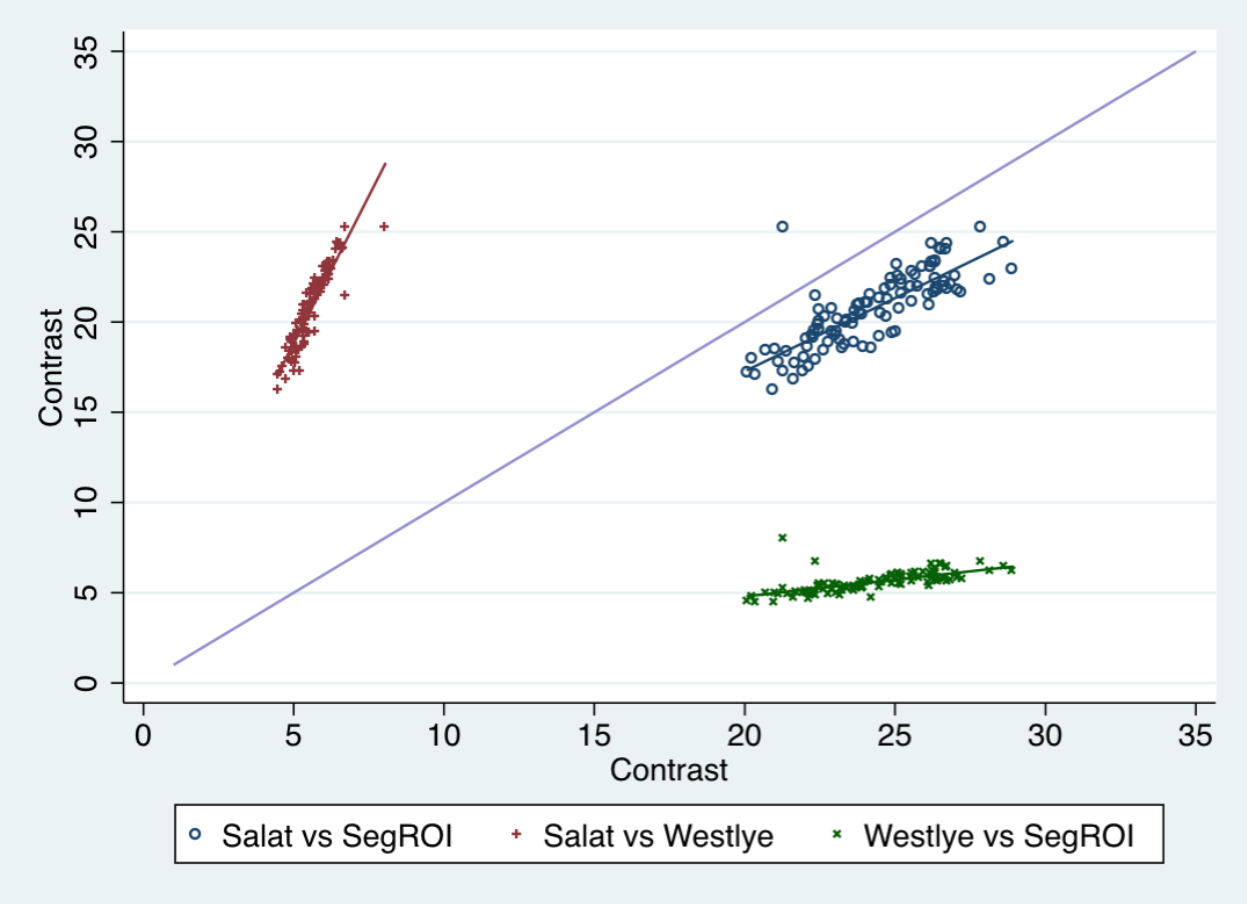

Figure 7. Scatter plot of agreement for each pairwise combination of measurements with line of equity overlaid on graph.

Table 11. Correlations with 95% Confidence intervals for each region and pairing of methods.

Region		correlation	95%	limits	METHOD1 = b(se)*METHOD2 + Constant(se)
Frontal					
Salat	Westlye	0.915***	0.877	0.942	Salat = 3.2660***(0.1440)Westlye + 2.4832**(0.8085)
Salat	SegROI	0.806***	0.725	0.865	Salat = 0.8103***(0.0595)SegROI + 1.0650(1.4496)
Westlye	SegROI	0.659***	0.533	0.757	Westlye = 0.1858***(0.0212)SegROI + 1.0782*(0.5157)
Insula					
Salat	Westlye	0.935***	0.905	0.956	Salat = 3.6205***(0.1376)Westlye + 1.5750*(0.6747)
Salat	SegROI	0.882***	0.83	0.919	Salat = 0.6642***(0.0355)SegROI + 3.8915***(0.8243)
Westlye	SegROI	0.76***	0.664	0.832	Westlye = 0.1479***(0.0126)SegROI + 1.4619***(0.2933)
Occipital					
Salat	Westlye	0.928***	0.895	0.951	Salat = 2.4803***(0.0998)Westlye + 5.2555***(0.5644)
Salat	SegROI	0.668***	0.545	0.764	Salat = 0.8355***(0.0930)SegROI + -0.7918(2.2226)
Westlye	SegROI	0.46***	0.292	0.601	Westlye = 0.2153***(0.0415)SegROI + 0.4596(0.9921)
Parietal					
Salat	Westlye	0.908***	0.867	0.937	Salat = 3.0846***(0.1420)Westlye + 2.9913***(0.8173)
Salat	SegROI	0.781***	0.692	0.847	Salat = 0.8188***(0.0654)SegROI + 1.0284(1.5727)
Westlye	SegROI	0.579***	0.434	0.695	Westlye = 0.1788***(0.0252)SegROI + 1.4407*(0.6048)
Temporal					
Salat	Westlye	0.95***	0.927	0.966	Salat = 3.5614***(0.1166)Westlye + 1.2365(0.6985)
Salat	SegROI	0.869***	0.811	0.91	Salat = 0.9918***(0.0565)SegROI + -2.1424(1.4065)
Westlye	SegROI	0.835***	0.764	0.885	Westlye = 0.2543***(0.0168)SegROI + -0.3485(0.4174)

*** p<0.05, ** p<0.01, *** p<0.001**

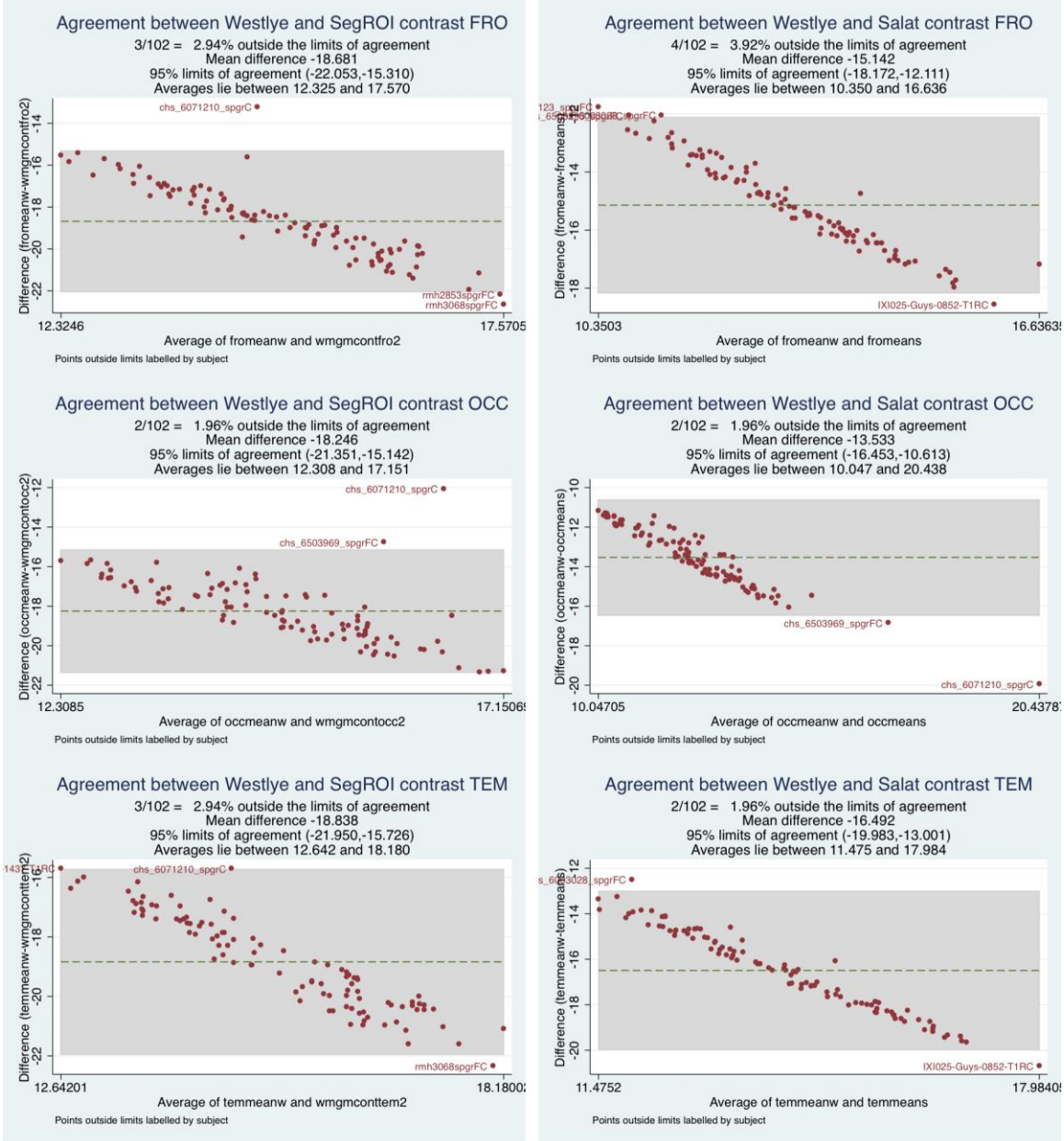

Figure 8. Bland-Altman plots of selected regional contrast between Westlye method and both SegROI and Salat methods.

Insula and Parietal not shown, but all show similar pattern of proportional bias.

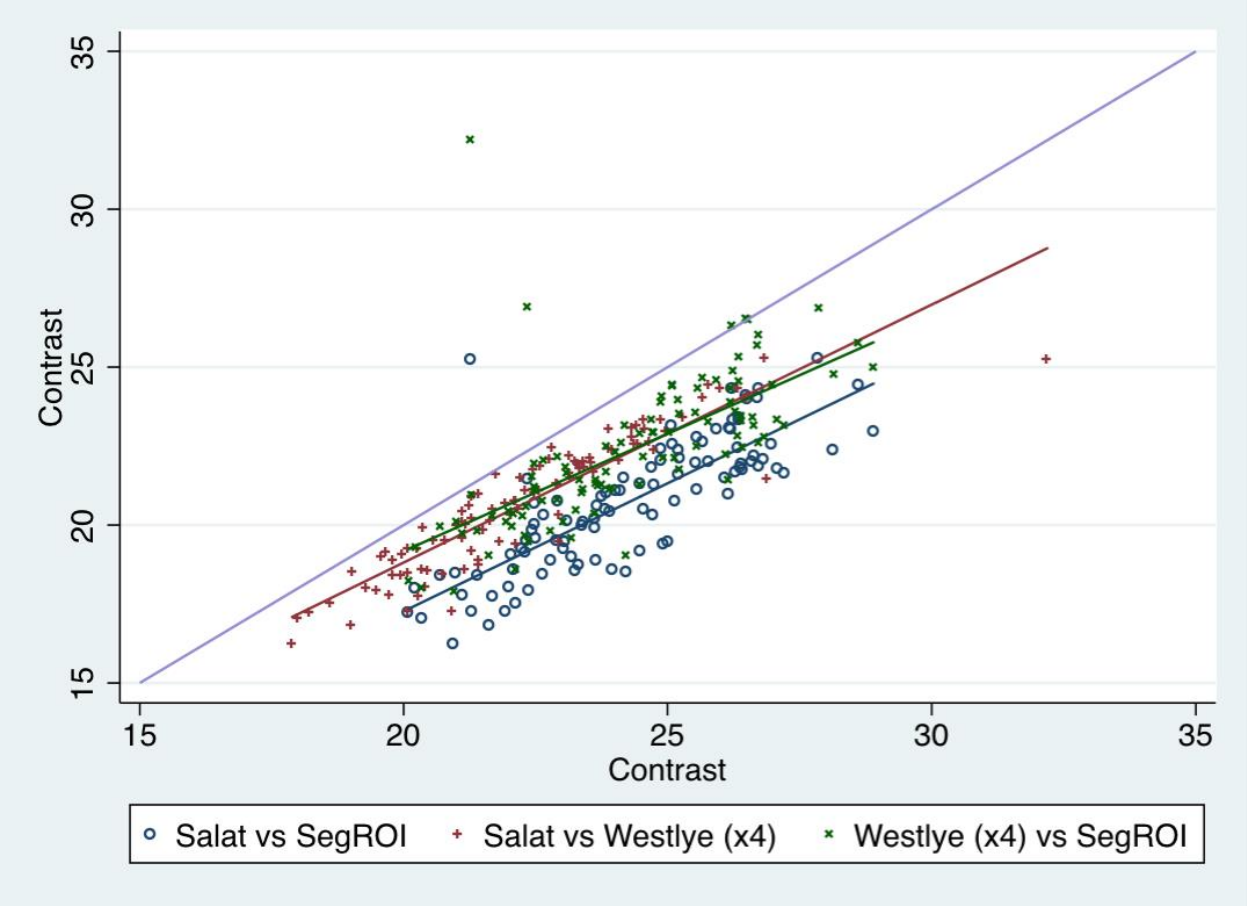

Figure 9. Scatter plot of agreement for each pairwise combination of measurements with line of equity overlaid on graph but with the Westlye method scaled up by a factor of 4.

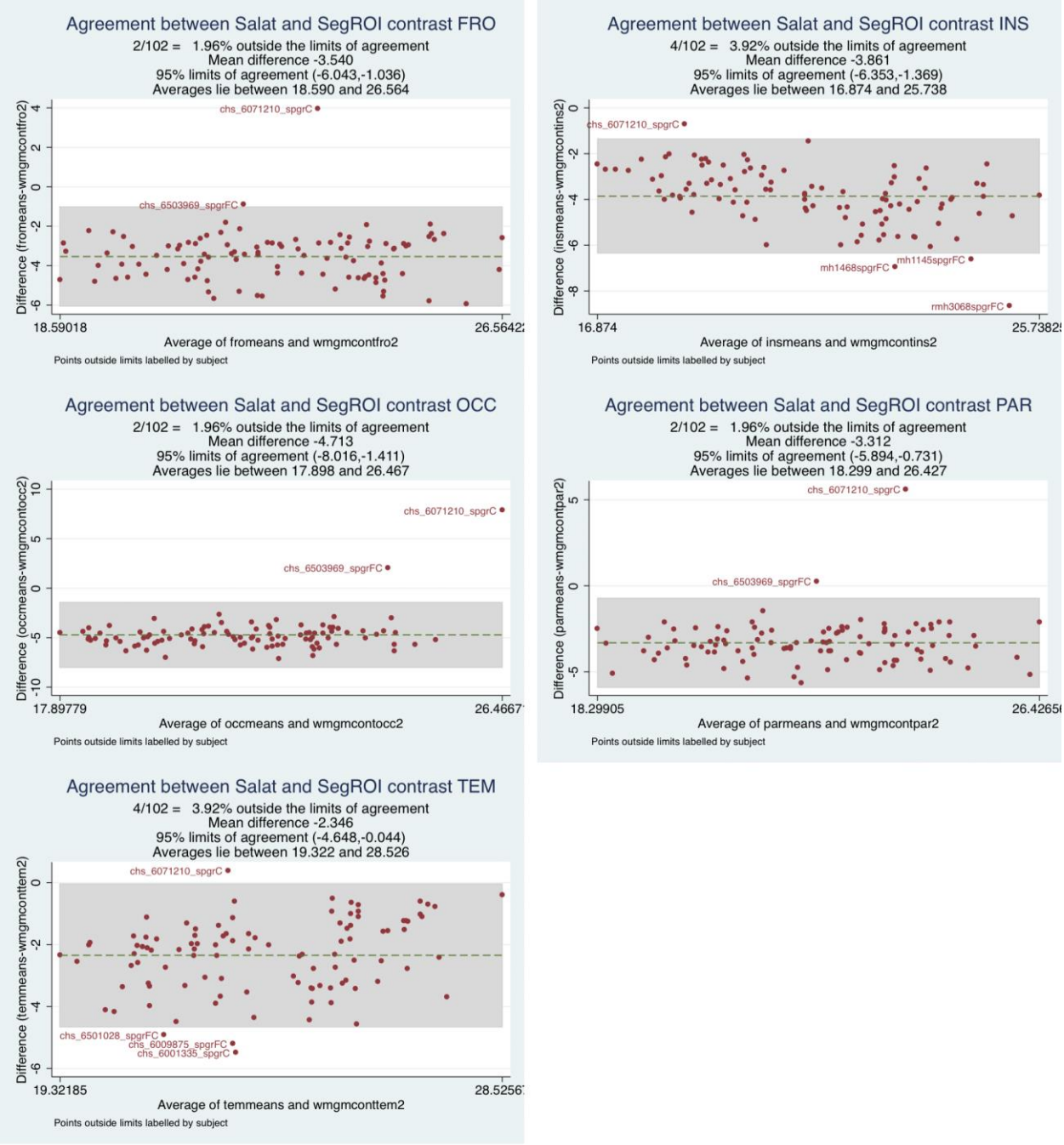

Figure 10. Bland-Altman plots of regional contrast between Salat method and SegROI method.

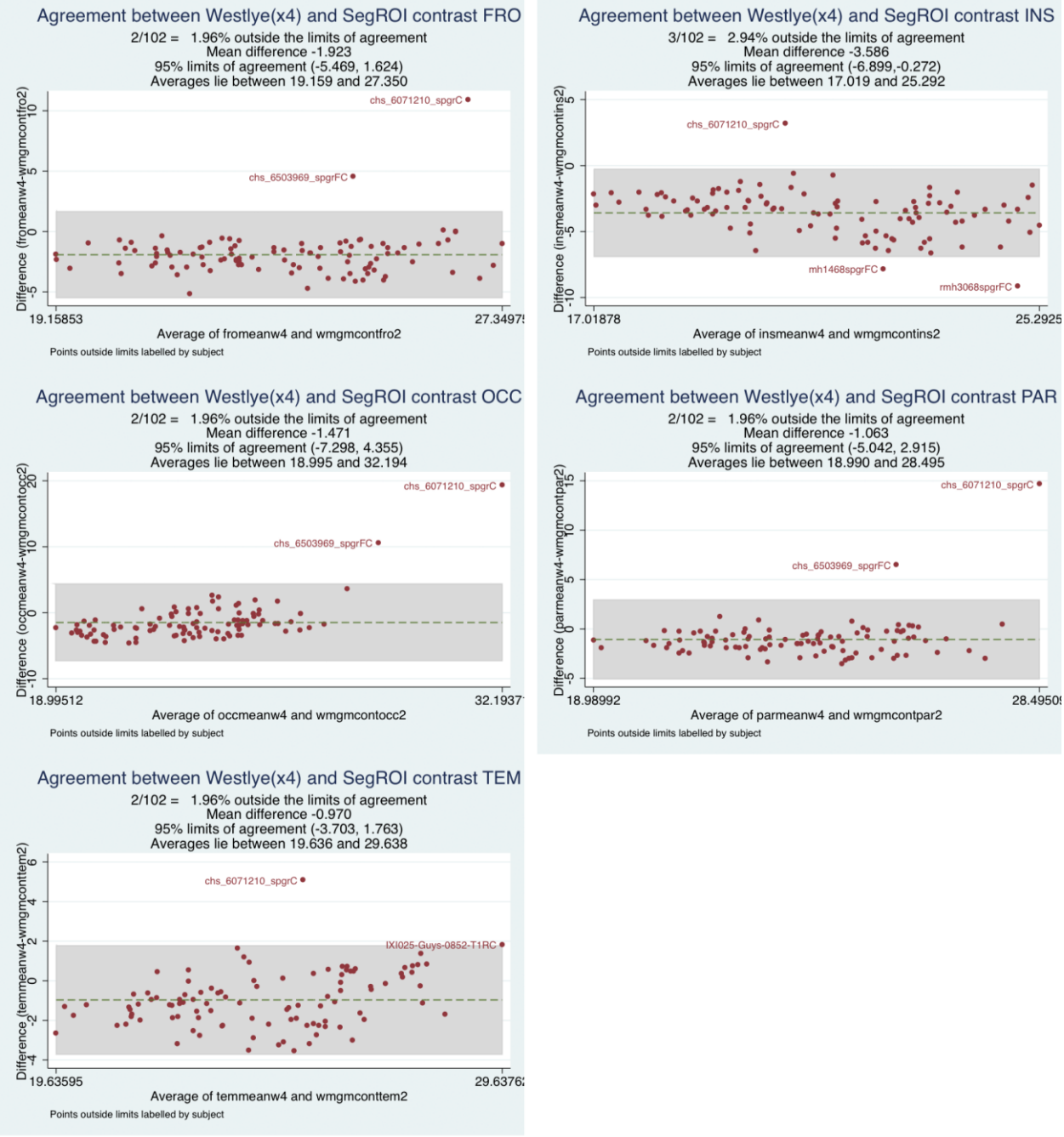

Figure 11. Bland-Altman plots of regional contrast between Westlye (scaled by 4) method and SegROI method.

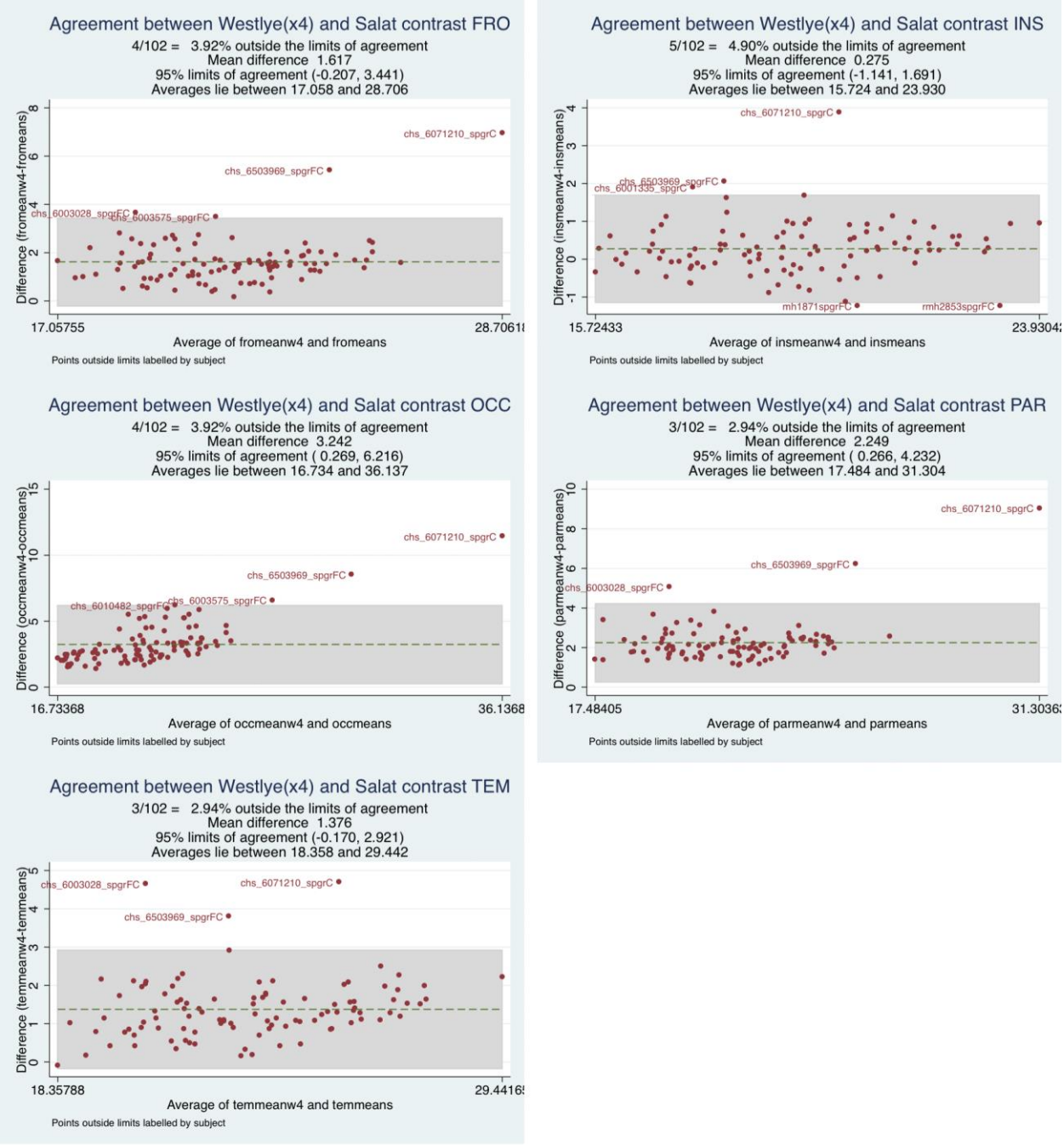

Figure 12. Bland-Altman plots of regional contrast between Westlye (scaled by 4) method and Salat method.

Table 12. Bias and limits of agreement for the three pairwise combinations using the scaled Westlye results.

	Bias	Standard Deviation	Percent error	Beyond limits
Salat vs SegROI				
Frontal	-3.504	1.277	20.63	2
Insula	-3.861	1.271	21.59	4
Occipital	-4.713	1.685	27.71	2
Parietal	-3.312	1.317	21.55	2
Temporal	-2.364	1.174	18.57	4
Westlye(x4) vs SegROI				
Frontal	-1.923	1.809	29.23	2
Insula	-3.586	1.691	28.70	3
Occipital	-1.471	2.973	48.88	2
Parietal	-1.063	2.030	33.21	2
Temporal	-0.97	1.394	22.05	2
Westlye(x4) vs Salat				
Frontal	1.617	0.931	17.60	4
Insula	0.275	0.722	14.73	5
Occipital	3.242	1.517	31.10	4
Parietal	2.249	1.012	19.21	3
Temporal	1.376	0.789	13.77	3

Table 13. Results of skewness and kurtosis tests for all pairwise combinations.

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	Adj chi2(2)	Prob>chi2
Salat vs SegROI					
Frontal	102	0.0000	0.0000	44.97	0.0000
Insula	102	0.0093	0.0362	9.64	0.0081
Occipital	102	0.0000	0.0000	.	0.0000
Parietal	102	0.0000	0.0000	70.70	0.0000
Temporal	102	0.1014	0.6880	2.92	0.2323
Westlye vs Salat					
Frontal	102	0.5144	0.0220	5.50	0.0639
Insula	102	0.4714	0.0259	5.35	0.0688
Occipital	102	0.0026	0.0067	13.47	0.0012
Parietal	102	0.4762	0.0901	3.47	0.1762
Temporal	102	0.7694	0.0033	7.83	0.0200
Westlye vs SegROI					
Frontal	102	0.3125	0.8514	1.08	0.5837
Insula	102	0.6992	0.0248	5.12	0.0773
Occipital	102	0.0321	0.0561	7.48	0.0238
Parietal	102	0.1572	0.1333	4.37	0.1124
Temporal	102	0.6792	0.0000	15.45	0.0004
Westlye(x4) vs Salat					
Frontal	102	0.0000	0.0000	58.72	0.0000
Insula	102	0.0000	0.0000	29.91	0.0000
Occipital	102	0.0000	0.0000	50.49	0.0000
Parietal	102	0.0000	0.0000	.	0.0000
Temporal	102	0.0000	0.0000	34.38	0.0000
Westlye(x4) vs SegROI					
Frontal	102	0.0000	0.0000	.	0.0000
Insula	102	0.8134	0.0021	8.46	0.0145
Occipital	102	0.0000	0.0000	.	0.0000
Parietal	102	0.0000	0.0000	.	0.0000
Temporal	102	0.0016	0.0038	14.84	0.0006

4.0 DISCUSSION

This research project examined four separate measures (signal intensity, contrast, CV and volume) with relation to age, gender and site. In addition we looked at the current popular method for extracting tissue contrast (FreeSurfer) and compared it to a more traditional ROI based sampling method. There are four main findings as a result of this research. First, it is problematic to make inferences about signal intensity without first normalizing the measure. Second, measures of contrast and CV are both normalized views of signal intensity and both show similar if inverse relationships with age. Third, volume changes seen here are in agreement with previously published results but show site related differences and if one normalizes volume results by ICV, site related differences appear to be minimized. Finally, the measures of contrast extraction put forth by current published FreeSurfer methods may not be optimal. More detailed analysis of the findings will be discussed below.

4.1 SIGNAL INTENSITIES

The raw signal intensities for all tissue segments decrease with age in two of the sites, while in the third site (Hammersmith Hospital, using a 3T scanner), all signal intensity measures increase with age. When a negative relationship with age is seen, WM has a steeper decline than GM,

which in turn has a steeper decline than CSF. But the question remains is it really safe to make assumptions on raw signal intensity values without attempting to normalize these values in some fashion? Likewise, should CSF signal intensity be decreasing at all?

The images from Hammersmith Hospital alone make a good argument against analyzing image signal intensities without first normalizing the values in some way. The fact that the gray-white contrast values follow the same trend as the other scanners after normalization is a sign that this may be an artificial problem. While it cannot be ruled out that the scanning sequence differs in some significant way between 1.5T and 3T, the more plausible explanation is that the images at Hammersmith Hospital are being scaled at reconstruction to use all of the available dynamic rage of the scanner, while at the other sites this does not appear to be the case. We are fairly certain that the images from the University of Pittsburgh were not scaled on the scanner, but further experiments and inquiry with the scanning manufacturer are required, especially before attempting to combine data sets across scanners.

The decrease in CSF signal intensity is also difficult to explain, but may be due to partial volume averaging. Partial volume averaging occurs when an image voxel encodes information for more than one type of tissue. For CSF voxels along the cortex, this neighboring tissue would be gray matter. For CSF voxels in the interior ventricles of the brain, the neighboring tissue is more likely white matter. These mixed tissue voxels can either be higher or lower than pure voxels of a single tissue type. For CSF voxels, mixing with neighboring tissue almost always increases the signal intensity. While there is little evidence of a biologic change in the composition of CSF with age, there are many published reports showing that the overall volume of CSF increases significantly with age[10, 11]. In the younger subjects, the overall CSF volume is much lower (~15% of inter-cranial volume) than the middle aged (~20% of ICV) and elderly

subjects (>20% of ICV). This makes sampling areas of pure CSF in the younger subject's images very difficult, and almost impossible to do in an automated fashion. As the number of pure CSF voxels increases with increased CSF volume and therefore age, the CSF signal intensity shows less partial volume averaging. This can be seen in Figure 13 showing the inverse relationship between volume of ventricles and signal intensity for subjects from Pitt and GH. If this is indeed the case, then using CSF signal intensity may not be a reasonable control region to use for normalizing results (by possibly calculating GM-CSF or WM-CSF contrast) especially in younger subjects. While similar arguments about partial volume averaging can be made for GM and WM, the strength of the relationship between signal intensity and volume was much lower for GM and not significant for WM.

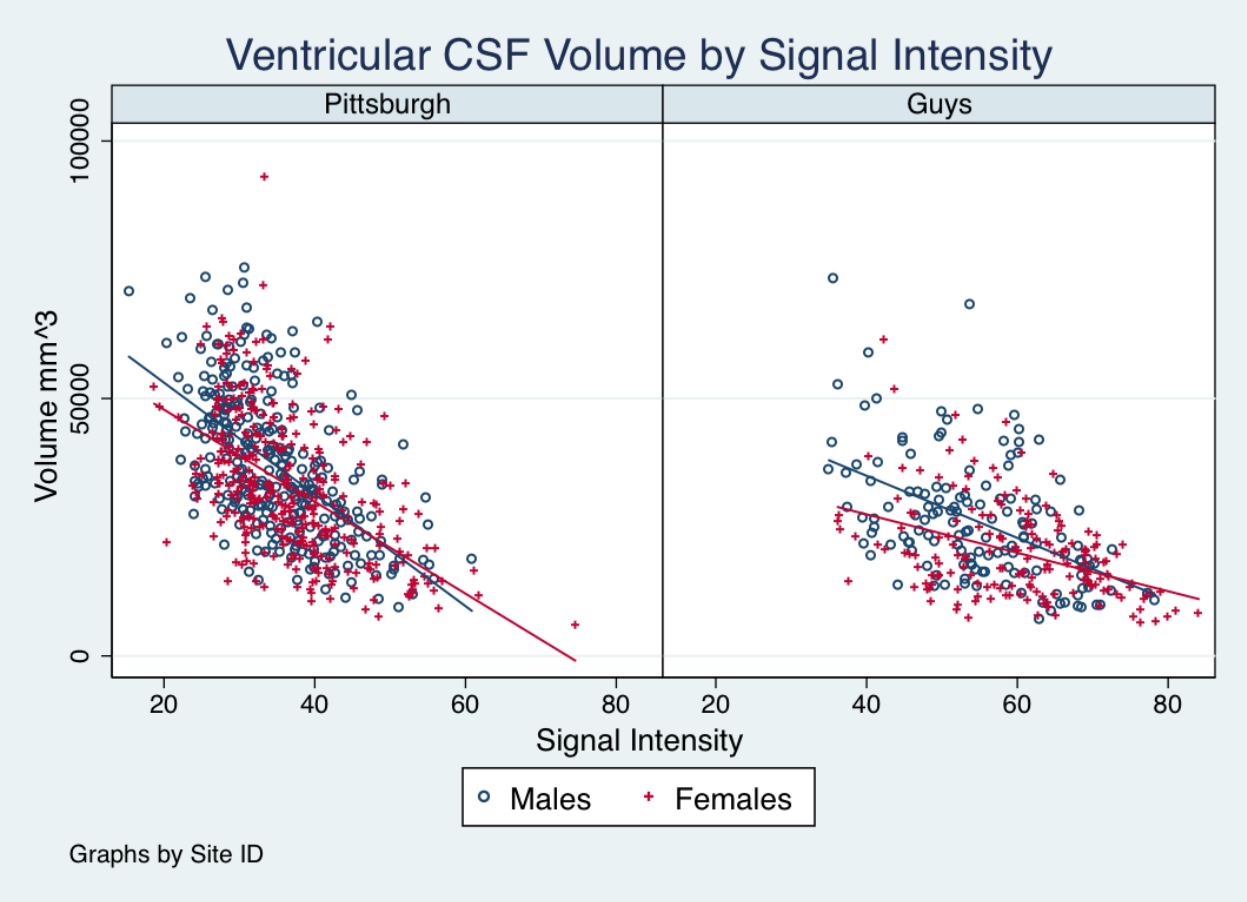

Figure 13. CSF Ventricular volume by Signal intensity.

In this research we see that signal intensities of T1 weighted MR images can be influenced by scanner site, as in the case of Hammersmith Hospital, as well as partial volume averaging in the case of the signal intensity measurements of CSF. Many additional variables including scanning sequence, receiver and transmission coils, magnet strength, and other software and hardware changes may also affect the quality of the signal intensities in some way and are the focus of large multi center studies such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) study. Therefore we can make no strong inferences about the relationship between signal intensity and age based on these results other than request that researchers carefully consider what it is they are looking at. It is imperative that the researcher understands how the images are being processed on the scanner, or counterintuitive inferences may be the result.

4.2 CONTRAST

The results of this study confirm the age related changes in gray-white contrast in standard structural T1 weighted MRI as seen in prior studies[23, 24, 26] while using completely novel datasets as well as using a different method of sampling signal intensities. While the recent studies on gray-white contrast have reported widespread contrast changes with several regions spared, namely occipital, and few gender related differences, our results appear to show this age related decline to be a more global phenomenon with the rate of decline differing based on gender and scanning site. The lowest rates of contrast decline are found in the Cerebellum and

Thalamus, areas that are not well reported in recent studies, while the highest rates of contrast decline are found in the Caudate, Insula and Parietal. The only region that did not show a decrease in contrast with age was the Putamen, a deep brain structure that is very difficult to segment into Gray Matter and White Matter and therefore a potentially unreliable region.

While these results confirm the results of prior studies, it is important to understand the difference in methodology. This methodology segments out the brain into tissue compartments before aggregating signal intensity values across an ROI, be it whole brain or a smaller sub-region like the Insula or Temporal lobe. This should produce a less noisy signal intensity measure, which may be why the within tissue compartment CVs are very low, but it will only be accurate in proportion to how accurately the segmentation algorithm works. Using an ROI approach also allowed us to better refine our models taking interactions between regression terms into account.

4.3 COEFFICIENT OF VARIATION

Global CV of GM is greater than that of WM. CV of CSF is greater than both. This result is expected as mean signal intensity (denominator) is decreasing WM>GM>CSF regardless of standard deviation.

Overall, the CVs are very low for within tissue segments. How much of this is due to the sampling being done after successful segmentation of the MRI is unknown. When using slightly different segmentation algorithms for SPM8 for a subset of data, VBM8 toolbox produces very similar and highly correlated mean and standard deviation values, while using SPM8's "new segment" produced similar means with even lower standard deviation measurements (data not

shown). Regardless, the significant differences seen in our models may be due to the large sample size and the small standard error estimates. It is not unreasonable to assume that given a different segmentation method, these differences may not be seen at all.

4.4 VOLUMES

Volume changes matched previously published reports with global decreases in GM volume related to age, gender and site[10, 11, 13, 50]. GM volume showed a steeper decline with age than WM volume. Males showed a steeper decline in both GM volume and WM volume than females. The site by age interaction may be due, in some part, to the differing age ranges of those populations, but precisely how much is unclear.

Normalizing each whole brain tissue compartment volume measure for ICV removed the gender by age and site by age interaction terms from the models in GM and CSF, but still showed small (~1% or less) but significant differences in the main effect of gender and site. Here the main effects of site can be traced back to the differing age range of the sites. Figure 14 shows that the increase with age of the CSF volume ratio to be highly comparable across sites. The ability to discern the one percent or less difference between the sites may be due to the large number of subjects in the study. A previous study on these same parameters (age, gender, and head size) showed strong relationships with volume, but did not look at across site relationships.[51]

As with previous studies[10, 11], the normalized WM volume appears to have more of a curvilinear path through adulthood, but the lack of younger subjects in the University of

Pittsburgh sample made it difficult to fit this model. Even the best model that included the base model with the addition of age^2 and age^2 by site interaction terms still had an R^2 of less than 0.14 whereas the standard model for normalized GM and CSF had R^2 of 0.71 and 0.65 respectively.

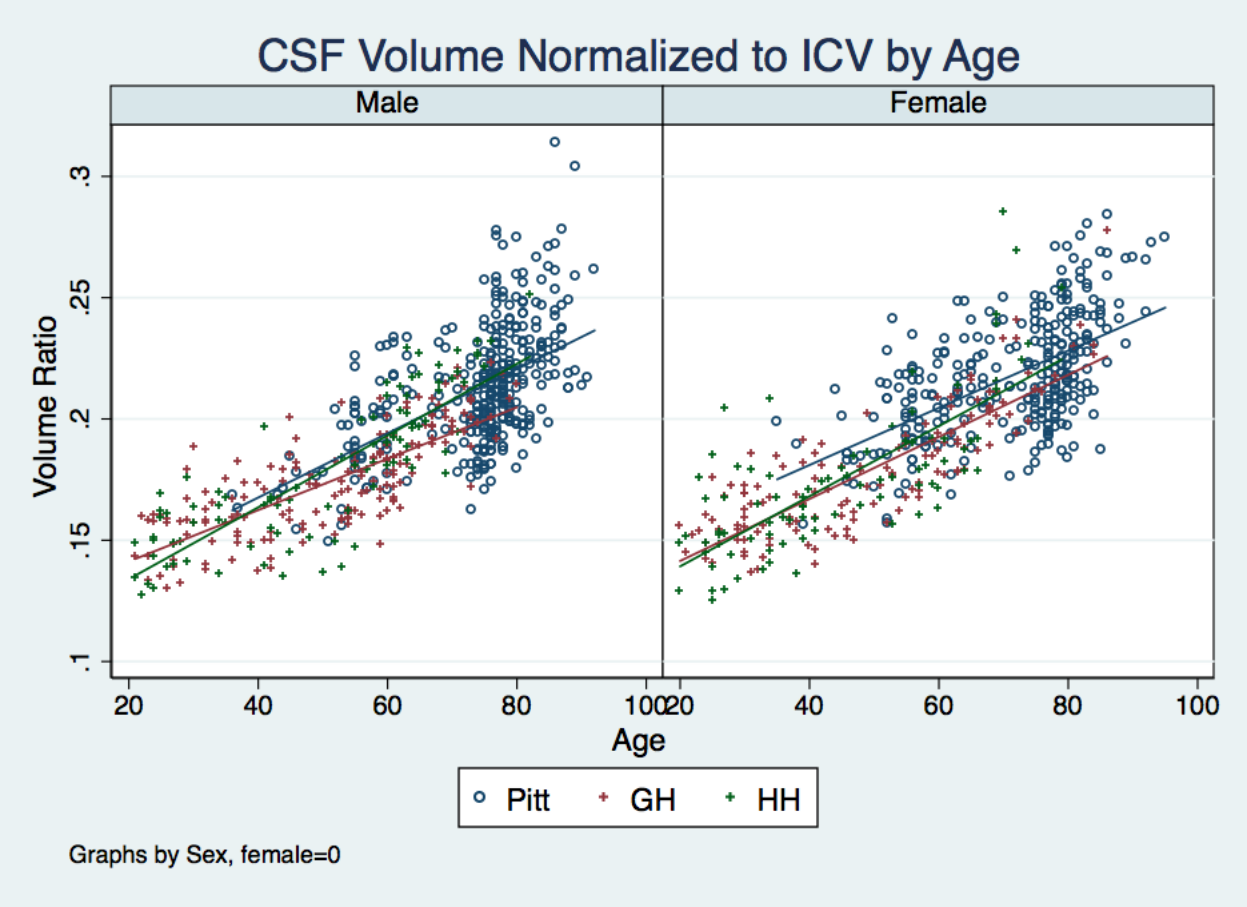

Figure 14. CSF volume normalized by ICV by age separated by gender and scanning site (University of Pittsburgh= blue, Guys Hospital=maroon, Hammersmith hospital=green).

4.5 AGREEMENT

No method produces the exact same results, even after scaling. The method of Westlye et al. produces the lowest measure of contrast, while the segmented ROI method produces the highest. The method of Salat et al. is closer to, but slightly lower than, the segmented ROI method. The two FreeSurfer methods are the most highly correlated. The segmented ROI method is on average more highly correlated with the method of Salat than the method of Westlye.

While the methods of Salat and Westlye are highly correlated, the results show a strong proportional bias. In essence, the values as calculated by Westlye's method (when calculated as a contrast measure, not a ratio as in the published results) are 3.5 to 4 times less than the values calculated by Salat's method, but the actual scaling value depends on the region being sampled. The large bias is most likely due to the larger distance between the sampling points in the method of Salat versus Westlye, while the regional variation may be due to the fixed versus variable method of sampling between the two methods. To understand the difference in these methods, it is important to understand how FreeSurfer finds the GM/WM boundary. First, note that the images used for this study are composed of voxels on the order of 1mm x 1mm x 1.5mm. FreeSurfer uses both signal intensity as well as 3-D intensity gradients to calculate the location of the GM/WM boundary. This boundary will be somewhere within a voxel of the image. While this method works well for finding accurate measurements of cortical thickness, sampling signal intensities too close to this boundary, which will be in interpolated data, will more than likely produce very underestimated values of the true tissue signal intensity.

The method of Westlye uses a fixed sample distance of 0.2mm from the gray-white border in both GM and WM. The two sample points are therefore 0.4mm apart. Salat on the other hand uses a fixed WM sample point 1.0mm from the border and a variable GM sample

point 35% of the distance from the border into GM. Depending on the thickness of the GM ribbon, the distance between sample points will vary accordingly but never be less than 1mm between the sample points. Regional variations of GM thickness may create regional differences in gray-white contrast. A more thorough comparison of all possible combinations of signal intensity sampling methods using FreeSurfer is needed.

Of greater concern for the FreeSurfer methods is the fact that both methods produce extreme values when sampling across a region. Table 14 shows maximum and minimum contrast values for the five regions samples using the FreeSurfer methods. Negative contrast values can occur when signal intensities for areas labeled as GM are higher than areas labeled as WM. Both FreeSurfer methods produce negative values, which may also be suppressing the true level of contrast. Also of note are some of the rather large maximum values specifically in the Occipital region. In order to obtain a contrast value of 200, the value of GM must be zero. These extreme values, both positive and negative, may present a problem when performing voxel by voxel statistical tests on the contrast maps in FreeSurfer.

Sampling too close to the GM/WM boundary produces very low contrast estimates in the case of the Westlye method, while both methods produce extreme values across a region. In spite of the fact that the Salat method samples GM and WM signal intensity a greater distance apart than Westlye, this measure of contrast is also consistently lower compared to the contrast calculated with the segmentation ROI method.

One subject from the Cardiovascular Health Study in Pittsburgh, a 72yo female, appears to be an outlier in the graphs. This subject's MRI was very noisy with possible motion artifacts, but the quality of the image was not so poor as to warrant excluding from the study. All three

methods produced results that differed significantly from the mean difference as seen on the Bland Altman plots.

Table 14. Minimum and maximum contrast values for the Freesurfer methods.

	Mean Contrast	Std. Dev.	Min Contrast	Max Contrast
Salat				
Minimum				
Frontal	-39.370	21.024	-115.968	-5.188
Insula	-16.911	12.092	-66.756	4.617
Occipital	-28.565	25.412	-116.396	-1.779
Parietal	-27.642	26.265	-123.090	-1.208
Temporal	-28.662	20.659	-126.559	1.016
Maximum				
Frontal	65.017	16.812	39.705	112.304
Insula	47.196	13.603	31.095	101.056
Occipital	56.160	22.943	34.938	200.000
Parietal	53.335	16.178	35.761	119.789
Temporal	60.861	18.733	38.412	146.105
Westlye				
Minimum				
Frontal	-11.326	6.040	-34.580	-3.096
Insula	-5.807	4.080	-25.721	-0.376
Occipital	-8.244	6.793	-29.691	-1.685
Parietal	-8.348	7.642	-31.422	-1.099
Temporal	-7.829	5.218	-34.436	-1.439
Maximum				
Frontal	26.468	8.661	14.297	48.933
Insula	15.680	5.691	8.588	34.774
Occipital	23.420	26.854	11.153	200.000
Parietal	18.888	9.240	10.658	66.720
Temporal	25.984	13.789	11.806	101.517

Values represent the Min or Max contrast calculated over the cortical region. Mean, SD, MIN and MAX are across all observations for selected region and method. The ability for the minimum contrast to be negative represents the GM signal intensity is greater than the WM signal intensity.

4.6 LIMITATIONS

In general, there are several caveats that must be mentioned. The signal intensities in an MRI may be influenced by changes in scanner software and hardware. In addition, how each manufacturer scales the data in an image acquisition is not always clear. Therefore, making any assumptions based on non-normalized signal intensity values may be dubious at best. While we are certain that the images from the University of Pittsburgh maintained the same hardware configuration throughout the duration of the studies, we are less certain of the effects that software changes may have contributed to our dataset, as well as what changes may have taken place during the acquisition phase of the open source datasets.

Partial volume averaging of voxels on the edge of the image segments may lead to lower estimates of contrast in a region. This effect is most pronounced in smaller regions or regions with small volumes of a given tissue segment. In an attempt to minimize this problem, we chose large "lobular" regions for this study thereby sampling larger volumes of any given tissue type. In addition we found that the younger subjects with the lowest CSF volumes had higher CSF signal intensities, which appear to be a sign of this specific partial volume effect.

The method of calculating contrast measures based on post segmented MRI, while not perfect, decreases the variability of the single point estimates being employed in FreeSurfer. Using a subsample of our data across all sites shows that the variation of the gray-white contrast within a region as calculated by FreeSurfer can be quite large going so far as to get a negative contrast score (GM signal value is larger than WM signal value). The method of sampling can also produce very different contrast measures.

5.0 CONCLUSION

The results of this research confirm the findings that there is a reduction in grey-white contrast with increased age. This appears to be repeatable across three separate sites despite the difference in slopes across those sites.

Even if the same manufacturer produces the MRI scanners, site matters greatly for all measures in this analysis. All analyses with the exception of volume normalized by inter-cranial volume showed significant age by site interactions, with some of the interactions being very dramatic. As a result, caution must be taken before combining data from separate sites. This may also lead to questions of repeatability of research projects between sites.

Questions remain as to the effect of contrast on imaging tools, tissue segmentation, cortical thickness measures, etc. Even though contrast and volume measures including cortical thickness appear to be statistically independent, further simulation studies that vary signal intensity across tissue groups should be done to assess the overall influence this basic measure has on downstream analyses.

APPENDIX A

MODEL KEY

Variable	Type	Coding
Gender	indicator variable	1.Gender: female=0, male=1
AgeC	continuous variable	centered to mean 62.55
Site	indicator variable	1.Site: GH=1
		2.Site: HH=1
		1.Site: 2.Site = 0 =Pitt
Gender#AgeC	interaction	1.Gender#AgeC: male*AgeC
Site#AgeC	interaction	1.Site#AgeC: GH*AgeC
		2.Site#AgeC: HH*AgeC

APPENDIX B

LIST OF COMMON ABBREVIATIONS

CSF	Cerebrospinal fluid
CV	Coefficient of Variation
ETL	Echo Train Length
GH	Guys Hospital London
GM	Gray Matter
HH	Hammersmith Hospital London
ICV	Inter-cranial volume
MNI	Montreal Neurologic Institute
MR(I)	Magnetic Resonance (Image/Imaging)
Pitt	University of Pittsburgh
RF	Radio Frequency
ROI	Region of interest
SD	Standard Deviation
SE	Standard Error
TE	Echo Time

TR Repetition Time

WM White Matter

BIBLIOGRAPHY

1. Bitar, R., et al., *MR pulse sequences: what every radiologist wants to know but is afraid to ask.* Radiographics, 2006. **26**(2): p. 513-37.
2. Jernigan, T.L., et al., *Cerebral structure on MRI, Part I: Localization of age-related changes.* Biol Psychiatry, 1991. **29**(1): p. 55-67.
3. Jernigan, T.L., et al., *Effects of age on tissues and regions of the cerebrum and cerebellum.* Neurobiol Aging, 2001. **22**(4): p. 581-94.
4. Lim, K.O., et al., *In vivo quantification of the limbic system using MRI: effects of normal aging.* Psychiatry Res, 1990. **35**(1): p. 15-26.
5. Pfefferbaum, A., et al., *A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood.* Arch Neurol, 1994. **51**(9): p. 874-87.
6. Raz, N., et al., *Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter.* Cereb Cortex, 1997. **7**(3): p. 268-82.
7. Salat, D.H., J.A. Kaye, and J.S. Janowsky, *Prefrontal gray and white matter volumes in healthy aging and Alzheimer disease.* Arch Neurol, 1999. **56**(3): p. 338-44.
8. Shear, P.K., et al., *Longitudinal volumetric computed tomographic analysis of regional brain changes in normal aging and Alzheimer's disease.* Arch Neurol, 1995. **52**(4): p. 392-402.

9. Sullivan, E.V., et al., *Age-related decline in MRI volumes of temporal lobe gray matter but not hippocampus.* Neurobiol Aging, 1995. **16**(4): p. 591-606.

10. Good, C.D., et al., *A voxel-based morphometric study of ageing in 465 normal adult human brains.* Neuroimage, 2001. **14**(1 Pt 1): p. 21-36.

11. Peelle, J.E., R. Cusack, and R.N. Henson, *Adjusting for global effects in voxel-based morphometry: gray matter decline in normal aging.* Neuroimage, 2012. **60**(2): p. 1503-16.

12. Sowell, E.R., et al., *Mapping cortical change across the human life span.* Nat Neurosci, 2003. **6**(3): p. 309-15.

13. Sowell, E.R., P.M. Thompson, and A.W. Toga, *Mapping changes in the human cortex throughout the span of life.* Neuroscientist, 2004. **10**(4): p. 372-92.

14. Fjell, A.M., et al., *High consistency of regional cortical thinning in aging across multiple samples.* Cereb Cortex, 2009. **19**(9): p. 2001-12.

15. Salat, D.H., et al., *Thinning of the cerebral cortex in aging.* Cereb Cortex, 2004. **14**(7): p. 721-30.

16. Head, D., et al., *Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the Alzheimer type: evidence from diffusion tensor imaging.* Cereb Cortex, 2004. **14**(4): p. 410-23.

17. Madden, D.J., et al., *Diffusion tensor imaging of adult age differences in cerebral white matter: relation to response time.* Neuroimage, 2004. **21**(3): p. 1174-81.

18. Pfefferbaum, A., E. Adalsteinsson, and E.V. Sullivan, *Frontal circuitry degradation marks healthy adult aging: Evidence from diffusion tensor imaging.* Neuroimage, 2005. **26**(3): p. 891-9.

19. Pfefferbaum, A. and E.V. Sullivan, *Increased brain white matter diffusivity in normal adult aging: relationship to anisotropy and partial voluming.* Magn Reson Med, 2003. **49**(5): p. 953-61.

20. Pfefferbaum, A., et al., *Age-related decline in brain white matter anisotropy measured with spatially corrected echo-planar diffusion tensor imaging.* Magn Reson Med, 2000. **44**(2): p. 259-68.

21. Salat, D.H., et al., *Age-related alterations in white matter microstructure measured by diffusion tensor imaging.* Neurobiol Aging, 2005. **26**(8): p. 1215-27.

22. Salat, D.H., et al., *Age-related changes in prefrontal white matter measured by diffusion tensor imaging.* Ann N Y Acad Sci, 2005. **1064**: p. 37-49.

23. Salat, D.H., et al., *Hippocampal degeneration is associated with temporal and limbic gray matter/white matter tissue contrast in Alzheimer's disease.* Neuroimage, 2011. **54**(3): p. 1795-802.

24. Salat, D.H., et al., *Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast.* Neuroimage, 2009. **48**(1): p. 21-8.

25. Westlye, L.T., et al., *Differentiating maturational and aging-related changes of the cerebral cortex by use of thickness and signal intensity.* Neuroimage, 2010. **52**(1): p. 172-85.

26. Westlye, L.T., et al., *Increased sensitivity to effects of normal aging and Alzheimer's disease on cortical thickness by adjustment for local variability in gray/white contrast: a multi-sample MRI study.* Neuroimage, 2009. **47**(4): p. 1545-57.

27. Kruggel, F., J. Turner, and L.T. Muftuler, *Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort.* Neuroimage, 2010. **49**(3): p. 2123-33.

28. Grydeland, H., et al., *Improved prediction of Alzheimer's disease with longitudinal white matter/gray matter contrast changes.* Hum Brain Mapp, 2012.

29. Cho, S., et al., *Establishing norms for age-related changes in proton T1 of human brain tissue in vivo.* Magn Reson Imaging, 1997. **15**(10): p. 1133-43.

30. Ogg, R.J. and R.G. Steen, *Age-related changes in brain T1 are correlated with iron concentration.* Magn Reson Med, 1998. **40**(5): p. 749-53.

31. Raz, N., D. Millman, and G. Sarpel, *Cerebral correlates of cognitive aging: Gray-white-matter differentiation in the medial temporal lobes, and fluid versus crystallized abilities.* Psychobiology, 1990. **18**(4): p. 475-481.

32. Steen, R.G., S.A. Gronemeyer, and J.S. Taylor, *Age-related changes in proton T1 values of normal human brain.* J Magn Reson Imaging, 1995. **5**(1): p. 43-8.

33. Piguet, O., et al., *White matter loss in healthy ageing: a postmortem analysis.* Neurobiol Aging, 2009. **30**(8): p. 1288-95.

34. Tang, Y., et al., *Age-induced white matter changes in the human brain: a stereological investigation.* Neurobiol Aging, 1997. **18**(6): p. 609-15.

35. Smith, S.M., et al., *Advances in functional and structural MR image analysis and implementation as FSL.* Neuroimage, 2004. **23 Suppl 1**: p. S208-19.

36. Lopez, O.L., et al., *Evaluation of dementia in the cardiovascular health cognition study.* Neuroepidemiology, 2003. **22**(1): p. 1-12.

37. Jennings, J.R., et al., *Brain imaging findings predict blood pressure response to pharmacological treatment.* Hypertension, 2008. **52**(6): p. 1113-9.

38. Jennings, J.R., et al., *Reduced cerebral blood flow response and compensation among patients with untreated hypertension.* Neurology, 2005. **64**(8): p. 1358-65.

39. Michelson, A., *Studies in Optics*1927: University of Chicago Press.

40. Dale, A.M., B. Fischl, and M.I. Sereno, *Cortical surface-based analysis. I. Segmentation and surface reconstruction.* Neuroimage, 1999. **9**(2): p. 179-94.

41. Fischl, B. and A.M. Dale, *Measuring the thickness of the human cerebral cortex from magnetic resonance images.* Proc Natl Acad Sci U S A, 2000. **97**(20): p. 11050-5.

42. Fischl, B., et al., *Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain.* Neuron, 2002. **33**(3): p. 341-55.

43. Fischl, B., et al., *Sequence-independent segmentation of magnetic resonance images.* Neuroimage, 2004. **23 Suppl 1**: p. S69-84.

44. Fischl, B., M.I. Sereno, and A.M. Dale, *Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system.* Neuroimage, 1999. **9**(2): p. 195-207.

45. Fischl, B., et al., *High-resolution intersubject averaging and a coordinate system for the cortical surface.* Hum Brain Mapp, 1999. **8**(4): p. 272-84.

46. Fischl, B., et al., *Automatically parcellating the human cerebral cortex.* Cereb Cortex, 2004. **14**(1): p. 11-22.

47. Segonne, F., et al., *A hybrid approach to the skull stripping problem in MRI.* Neuroimage, 2004. **22**(3): p. 1060-75.

48. Segonne, F., J. Pacheco, and B. Fischl, *Geometrically accurate topology-correction of cortical surfaces using nonseparating loops.* IEEE Trans Med Imaging, 2007. **26**(4): p. 518-29.

49. Bland, J.M. and D.G. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet, 1986. **1**(8476): p. 307-10.

50. Raz, N. and K.M. Rodrigue, *Differential aging of the brain: patterns, cognitive correlates and modifiers.* Neurosci Biobehav Rev, 2006. **30**(6): p. 730-48.

51. Barnes, J., et al., *Head size, age and gender adjustment in MRI studies: a necessary nuisance?* Neuroimage, 2010. **53**(4): p. 1244-55.