# RELIABLY MEASURING BRAIN WHITE MATTER HYPERINTENSITIES IN T2 FLAIR MRI

By

STEPHEN D. TOWLER

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2011

To my wife

ACKNOWLEDGMENTS

I thank my mentors and collaborators for their time and energy. I am especially grateful for the data collection performed prior to the present work, the time spent learning and applying the technique presented within, and assistance with debugging. This work was also made possible by ImageJ its open community of users and contributors.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS ..... 4

LIST OF TABLES ..... 7

LIST OF FIGURES ..... 8

LIST OF ABBREVIATIONS ..... 9

ABSTRACT ..... 10

CHAPTER

1 INTRODUCTION ..... 12

Methods for Measuring Severity of White Matter Hyperintensities ..... 12
Subjective Visual Ratings: Fast and Popular, but Low Fidelity ..... 13
Merits and limitations ..... 15
Use as a standard in the present work ..... 16
Manual Lesion Mapping: Implemented and Analyzed in the Present Work ..... 17
Merit: detailed evaluation of repeated-measure reliability ..... 18
Merit: provides a physical tissue volume ..... 20
Merit: allows integration with other spatial maps ..... 22
Merit: resulting maps can be used to train people and computers ..... 24
Automated Lesion Mapping: an Evolving Alternative ..... 25
Study Rationale ..... 26
Aims and Hypotheses ..... 26

2 METHODS ..... 28

Participants – Brain Magnetic Resonance Images ..... 28
Existing Subjective Ratings of White Matter Hyperintensities ..... 30
Novel Manual Segmentation of White Matter Hyperintensities ..... 30
Statistical Evaluations of Reliability ..... 32
Inter- and Intrarater Reliability of White Matter Hyperintensity Segmentation ..... 32
Convergent Validity of Methods ..... 33

3 RESULTS ..... 34

Aim 1: Inter- and Intrarater Reliability of White Matter Hyperintensity Segmentation ..... 34
Aim 2: Convergent Validity of Methods ..... 35

4 DISCUSSION ..... 37

Clinical Relevance of White Matter Hyperintensity Severity ..... 38
White Matter Hyperintensity is Associated with Clinical and Cognitive Deficits ..... 38
Severity matters ..... 39
Location matters ..... 41
Progression matters ..... 43
White Matter Hyperintenisty is Associated with Advancing Age and Cardiovascular Risk Factors ..... 44
Ongoing Extension ..... 45
Limitations ..... 46

LIST OF REFERENCES ..... 49

BIOGRAPHICAL SKETCH ..... 57

LIST OF TABLES

Table page

2-1 Demographics of brain magnetic resonance image participants ..... 29

3-1 Manual three-dimensional digital white matter hyperintensity segmentations are reliable within and between raters. ..... 34

3-2 White matter hyperintensity map volumes are reliable with previously assigned Junque scores. ..... 35

3-3 Summary statistics of reliability of Junque scores and white matter hyperintensity map volumes ..... 35

LIST OF FIGURES

Figure page

2-1 Creation of a three-dimensional white matter hyperintensity binary mask for each rating. ..... 32

3-1 Scatter plot of Rater 1 white matter hyperintensity volumes and previously acquired Junque scores. ..... 36

3-2 Scatter plot of Rater 1 white matter hyperintensity volume ranks and ranks of previously acquired Junque scores. ..... 36

## LIST OF ABBREVIATIONS

| | |
|---|---|
| 3D | three-dimensional |
| AD | Alzheimer’s disease |
| AIREN | Association Internationale pour la Recherché et l'Enseignement en Neurosciences |
| ASL | arterial spin labeling |
| DSC | Dice similarity coefficient |
| FLAIR | fluid-attenuated inversion-recovery |
| ICC | intra-class correlation |
| LA | leukoaraiosis |
| mm3 | cubic millimeter |
| MR | magnetic resonance |
| MRI | magnetic resonance imaging |
| NAWM | normal-appearing white matter |
| NINDS | National Institute of Neurological Disorders and Stroke |
| T1 | T1 – a magnetic resonance tissue property |
| T2 | T2 – a magnetic resonance tissue property |
| UMDNJ | University of Medicine and Dentistry of New Jersey |
| UBO | unidentified bright object |
| VaD | vascular dementia |
| WM | white matter |
| WMA | white matter abnormalities (or alterations) |
| WMH | white matter hyperintensities |
| WML | white matter lesions |

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

RELIABLY MEASURING BRAIN WHITE MATTER HYPERINTENSITIES IN T2 FLAIR MRI

By

Stephen D. Towler

August 2011

Chair: Bruce Crosson
Major: Psychology

White matter hyperintensity (WMH) seen on T2 fluid-attenuated inversion recovery (FLAIR) brain magnetic resonance imaging (MRI) is a neuroimaging sign observed in normal and abnormal aging. Severity of WMH is an important diagnostic criterion in vascular dementia, however severity is typically quantified using subjective visual rating scales, which have a number of significant limitations including poorly characterized reliability and low-fidelity output that communicates little detail about raters' judgments. In the present work the author aims to 1) develop a novel and efficient three-dimensional (3D) manual WMH mapping technique with high inter- and intrarater reliability, and 2) evaluate the convergent validity of this novel technique and an established WMH visual rating scale. *Methods*: In a cross-sectional observational design, two expert raters applied the novel WMH mapping technique to 15 T2 FLAIR MRIs previously scored with a widely-used WMH visual rating scale, the Junque scale. Each rater mapped WMH in each MRI twice, blind to identity, and one rater mapped all 77 MRIs for which Junque scores were available. *Results*: Spatial overlap of 3D maps was high in all inter- and intrarater comparisons (Dice similarity coefficient grand mean = 0.84, SD = 0.12). Tissue volumes calculated from repeated WMH mappings were

reliable with each other (ICC = .96, $p < 0.001$), and were correlated with the previously obtained Junque visual ratings (all Spearman's rho > 0.75, $p<0.01$). CONCLUSION: A novel technique for manual WMH segmentation is reliable within and between raters and with a well-known subjective rating scale previously shown to correlate with clinical and cognitive measures. Benefits of 3D WMH mapping over visual rating scales are discussed.

# CHAPTER 1
# INTRODUCTION

## Methods for Measuring Severity of White Matter Hyperintensities

White matter hyperintensity (WMH) is a neuroimaging sign frequently seen on T2 fluid-attenuated inversion recovery (FLAIR) brain magnetic resonance images (MRI) of healthy and demented individuals (Figure 2-1). In these brain images, WMH is observed in the brain's white matter, and represents damage to this axon-packed tissue responsible for the transport of neuronal impulses from gray matter neuron cell bodies to their eventual destinations. Evidence from post-mortem tissue examination (Young, Halliday, & Kril, 2008) suggests that WMH is a marker for white matter damage caused by the breakdown of intracranial vasculature (Pantoni & Garcia, 1997; Skoog, Marcusson, & Blennow, 1998; Vermeer et al., 2002). Though WMH is a common incidental finding in normally functioning individuals, WMH severity has been shown to be related to clinical and cognitive decline in patients (Libon, Price, Davis Garrett, & Giovannetti, 2004), and is currently used in the diagnosis of vascular dementia (Cosentino, Jefferson, M. Carey, Price, Davis Garrett, Swenson, et al., 2004a). In the present work the term "white matter hyperintensities" (or "WMH") is used to refer to this neuroimaging sign, but a number of synonymous terms appear throughout the last 40 years of literature, including age-related white matter changes, Binswanger's disease, brain rust, deep white matter changes, diseased white matter, hyperintense white matter, leukoaraiosis (LA), leukoencepathology, leukomalacia, T1 hypointensities, senile leukoencephalopathy, T2 hyperintensities, white matter abnormalities (WMA), white matter alterations (WMA), white matter changes, white matter high signal areas, white matter hypodensities, white matter lesions (WML), white matter pathology, white

matter rarefaction, white matter signal abnormalities (WMSA), and unidentified bright objects (UBO).

White matter hyperintensity on MRI is a biomarker that can be acquired quickly and non-invasively, requiring only five to fifteen minutes of MRI scanner time per whole-brain T2 FLAIR image (Rydberg et al., 1994), and no ingestion or injection of a contrast agent. However, to explore the relationships among WMH severity, demographic variables, cognitive deficits, and underlying pathophysiology, investigators must *quantify* the severity of observed WMH with a method that is valid, reliable, and practical. Much of the literature describing relationships between WMH and cognitive variables is based on semi-quantitative subjective visual ratings of WMH severity. In contrast, the present work describes a novel and efficient technique for manual digital three-dimensional (3D) mapping of WMH tissue. Each of these two classes of techniques has merits and limitations, described below.

**Subjective Visual Ratings: Fast and Popular, but Low Fidelity**

Until recently, investigators characterizing the severity of WMH primarily used subjective visual rating scales to provide semi-quantitative ratings of severity. These subjective visual rating scales provide a variety of methods for characterizing the proportion of an individual's brain tissue that appears as WMH on MRI images. Commonly these subjective visual ratings are performed by an expert rater who views an individual's single-time-point (as opposed to longitudinal) T2 FLAIR MRI brain images on a computer screen, and applies a scoring rubric to produce a numerical rating that represents the rater's judgment of the severity of WMH appearing on the MRI, with higher numerical ratings representing more severe WMH.

Though the present work employs just one visual rating scale for comparison with 3D WMH mapping, the Junque scale (Junque et al., 1990) (subjective visual rating range of 0 to 40), a number of other subjective WMH visual rating scales have been used to characterize WMH, including the Fazekas (Fazekas, Chawluk, Alavi, Hurtig, & Zimmerman, 1987) and Schmidt (R. Schmidt et al., 1992) scale (range of 0 to 3), the Scheltens scale (Scheltens et al., 1993) (range of 0 to 84), the Age-Related White Matter Changes (Wahlund et al., 2001) (range of 0 to 30), and the Manolio scale (Manolio et al., 1994) (range 0 to 9). These scales differ not only in range of final scores, but also in the anatomy attended to per each method's rating guidelines. For example, periventricular and deep-white-matter lesions are scored separately in scales of Fazekas and Schmidt, the scale of Scheltens, and the scale of Junque, but not the scale of Manolio. Even among the methods that score periventricular lesions separately, there are differences in how those periventricular lesions are characterized, with the Junque score dependent upon the rater's judgment of percentage of hyperintense tissue (< 25%, 25% to 50%, 50% to 75%, or > 75%), the Fazekas and Schmidt score dependent upon broad qualitative categories ("pencil-thin lining", "smooth halo", "irregular…extending into deep white matter"), and the Scheltens score dependent upon linear dimensions of observed WMH (<=5 mm, or > 5 mm). These scales also differ in a number of measured psychometric properties (Libon et al., 1998), but have all been used to report relationships between WMH severity and clinical, cognitive, and demographic variables. Unfortunately, comparisons among these visual rating scales have revealed inconsistent between-scale reliability and also variable sensitivity to

relationships between rated WMH severity and clinical, cognitive, and demographic variables (Kapeller et al., 2003; van Straaten et al., 2006).

**Merits and limitations**

Despite these questions of between-scale reliability, subjective visual rating scales do offer a number of methodological benefits. One is that subjective visual ratings are robust: humans build robust mental prototypes of what known patterns should look like, and can be easily trained to ignore irrelevant visual information. Our visual system is not easily confused by visual features that might confound automated computer algorithms, such as differences among MRI scanners or their settings, image artifacts from small head movements, or artifact generated by fluctuations in the functioning of MRI equipment. Another methodological advantage is that visual rating scales are fast: once trained, an expert rater can assign a subjective rating to a brain's WMH severity much faster than they can manually trace and quantify WMH on MRI. This efficiency is especially important given that the experts performing these ratings are typically clinical or research faculty with limited time. Finally, subjective visual rating is convenient and low-capital: it can be performed with any MRI viewing software that allows the rater to control image brightness and contrast. This lowers equipment and software barriers that might otherwise deter an investigator with valuable WMH images from analyzing them.

Despite these merits, subjective visual rating scales suffer from a number of methodological limitations sufficiently problematic to motivate the current work. Most egregious is the low fidelity output provided by subjective visual rating scales: everything an expert rater observes about a brain's WMH tissue is reduced to a single score or, at best, a few scores per hemisphere. This fails to communicate valuable judgments about lesion characteristics, such as anatomical locations, boundaries, and

MR signal contrast with normal-appearing tissue. This lack of recorded detail can also make it difficult and time consuming to train new raters: completed subjective visual ratings do not implicitly produce visual records that could be used as training sets, so experienced raters are relegated to co-rating many time-consuming exemplars with new trainees. Finally, there are limits to the detail with which the reliability of repeated subjective visual ratings can be investigated. Two independent raters might arrive at identical visual rating scores for a given brain, but may have done so by chance, having unknowingly attended to substantially different anatomical areas or tissue characteristics.

In addition to these methodological limitations, there is some evidence that subjective visual rating scales may not be as sensitive as WMH volumes calculated from WMH mapping techniques. A 2006 comparison of the Scheltens visual rating scale and an automated method for segmenting WMH and calculating WMH volume (den Heuvel et al., 2006) found that the baseline correlation between WMH severity and age was weaker when WMH severity was measured with the Scheltens visual scale than with the volumetric method. Longitudinal analysis of follow-up images (delay mean, SD = 33, 1.4 months) revealed that the correlation between WMH progression and age was twice as high for WMH volume than WMH subjective visual rating. Sensitivity to longitudinal changes is especially important in light of recent findings suggesting that persistent cognitive impairment may be predicted by the *progression* of WMH severity but not baseline severity (Silbert, Howieson, Dodge, & Kaye, 2009).

**Use as a standard in the present work**

Though they suffer from the limitations described above, previously obtained Junque visual ratings of WMH severity are used in the present work for comparison with

a novel 3D WMH manual mapping method. The Junque scale offers a number of advantages over other WMH subjective visual rating scales: Junque scores are interval data that do not suffer from restricted range (Cosentino, Jefferson, M. Carey, Price, Davis Garrett, Swenson, et al., 2004a), and have psychometric properties shown to be appropriate for both dependent and explanatory variables (Libon et al., 1998). As of this writing 199 articles have cited the Junque scale, and it has been proposed for use as the sole neuroradiological criterion in diagnosis of vascular dementia (Cosentino, Jefferson, M. Carey, Price, Davis Garrett, Swenson, et al., 2004a). A number of recent studies have used it to investigate links between WMH and motor or cognitive phenomena, including working memory deficits in dementia (Lamar et al., 2007; Lamar, Catani, Price, Heilman, & Libon, 2008), syntactic comprehension in dementia (Giovannetti et al., 2008), dementia symptom profiles (Libon et al., 2008; Price, Jefferson, Merino, Heilman, & Libon, 2005a), activities of daily living in dementia patients (Giovannetti, K. S. Schmidt, Gallo, Sestito, & Libon, 2006), clock drawing in dementia patients (Cosentino, Jefferson, Chute, Kaplan, & Libon, 2004b), and processing speed in aging adults (Junque et al., 1990). The specific Junque guidelines for scoring WMH severity are reproduced in the methods section below.

**Manual Lesion Mapping: Implemented and Analyzed in the Present Work**

The current work examines an alternative method of quantifying WMH: manual 3D lesion mapping, or “binary segmentation”, which does not suffer from the subjective visual rating scale limitations described above, and confers a number of methodological advantages, discussed below. In this technique an expert rater views each slice of a T2 FLAIR brain MRI and uses software to create a slice-by-slice map of points in space they endorse as WMH. In the implementation developed for the present work, this takes

a trained rater approximately 5 to 30 seconds per slice, or less than 15 minutes per brain for the 26-slice T2 FLAIR images analyzed here. These per-slice WMH maps are automatically concatenated in space to produce a perfect 3D record of the tissue the rater endorsed as WMH. Although this technique carries greater time and equipment burdens than subjective visual rating scales, it confers a number of methodological advantages simply by recording a rater's judgments as points in space.

**Merit: detailed evaluation of repeated-measure reliability**

The 3D lesion masks output during binary WMH segmentation facilitate detailed evaluation of agreement among repeated ratings. One form of evaluation is qualitative visual inspection of areas of non/overlap in repeated ratings, a procedure that is impossible with subjective visual rating scales but simple to perform for 3D tissue maps. This is useful during consensus conference or for individual raters, especially trainees, to directly visualize how WMH endorsements strategies or errors vary among repeated ratings of a given image. The spatial differences between two repeated ratings of the same brain become immediately obvious when the two lesion maps are displayed as translucent, overlapping, individually-colored layers.

Three-dimensional digital segmentation also offers investigators a precise method for *quantifying* the reliability of repeated measures: the proportion of spatial overlap between repeated ratings provides an index of agreement that can provide even more information than the Pearson product-moment correlation or intra-class correlation (ICC) of endorsed lesion volumes. Intra-class correlation is a useful method for analyzing reliability of repeated quantitative measures that meet ICC distributional assumptions, but when the evaluated ratings exist in two-dimensional (2D) or 3D space, measures of spatial overlap provide richer reliability data by indicating not only whether the

calculated areas (or volumes) of two ratings are similar, but also the degree to which the endorsed regions overlap in space. As an analogy, imagine training cartographers to manually draw lakeshores from new, experimental satellite data. The true measure of whether two repeated ratings are reliable with each other is the degree to which the resulting lake areas overlap each other in 2D space, not just the similarity in the number of square miles encompassed by each rating's lakeshore boundaries. Applied over a number of repeated ratings, ICC analysis might communicate that the square-millage of repeated ratings is similar, without communicating the degree to which those repeated ratings do or do not overlap in 2D space. The spatial overlap index chosen for the present work is the Dice similarity coefficient (DSC), calculated for pairs of repeated ratings. It is calculated as the overlap volume common to two ratings divided by one half the sum of their volumes. A DSC value of 1 represents perfect overlap, 0 represents no overlap, and values between 1 and 0 are interpreted like kappa, with > .7 indicating excellent agreement (Zijdenbos, Dawant, Margolin, & Palmer, 1994). This intuitive interpretation is why DSC was chosen from among a number of other reliability metrics that can be applied across space, including the Jaccard index (Jaccard, 1912) and the Tanimoto coefficiet (Crum, Camara, & Hill, 2006). The Sørensen similarity index (Sørensen, 1957) also appears in the literature but is computationally identical to the DSC.

Investigators wishing to thoroughly characterize the inter- and intrarater reliability of WMH observations should use 3D mapping techniques when setting, time, and equipment permit. The 3D spatial output allows for direct visualization of areas in which repeated ratings agree or disagree. Discrepancies among ratings can be displayed as

overlays on the actual brain MRI, providing a useful training tool and method for efficient assessment of error or disagreement. Also, by employing overlap measures such as the Dice similarity coefficient, the proportion of spatial agreement in repeated ratings can be quantified, providing more information than metrics such as intraclass correlation, which conveys the relationships among the magnitude of repeated ratings but not the degree of spatial overlap between repeated ratings.

**Merit: provides a physical tissue volume**

Three-dimensional WMH segmentation allows for the calculation of the physical volume occupied by WMH tissue, which can then be analyzed against other biological, cognitive, or clinical variables. Because it is a physical volume, it can also be expressed as a proportion of other physical volumes, such as total intracranial volume, total parenchyma volume, or total white matter volume. This permits parity with many subjective visual rating scales in which scores are based on the *proportion* of affected tissue, instead of just the magnitude.

Distributional qualities of output data limit the statistical approaches that can be applied to the data. Compared to visual rating scales, WMH volumes calculated from 3D mapping have distributions with superior statistical properties against which to examine clinical and cognitive variables: while subjective visual rating scales suffer from restricted range and a number of psychometric limitations (Libon et al., 1998), WMH volumes reflect the underlying physiological tissue volume to the degree that the MRI acquisition allows. This produces volumetric data with distributions that are readily addressable with parametric statistical techniques.

In addition to these methodological advantages, recent empirical evidence suggests that calculating WMH volumes provides superior severity data. One recent

cross-sectional study established that WMH volume is a more sensitive indicator of WMH severity than WMH subjective visual ratings (den Heuvel et al., 2006). Additionally, a recent longitudinal study highlights the importance of calculating WMH volumes by demonstrating that persistent cognitive impairment was predicted by the *progression* of total and periventricular WMH volumes, but not by baseline WMH volumes (Silbert et al., 2009). In longitudinal studies, changes in individuals' WMH can be measured as a difference in physical volume, providing a rate of WMH change expressed in unit of volume per unit of time. Furthermore these brain changes can also be localized by mapping the locations of WMH changes over time, creating a natural history of an individual brain's pathological changes. This permits the study of how specific regional neuropathological changes are linked to progressive neuropsychological decline.

Calculation of WMH volume is not unique to the manual 3D mapping technique presented in the current work. WMH tissue volumes can be similarly calculated from *automated* 3D mapping techniques, discussed later, and even by pixel-counting methods that do not produce verifiable 2D or 3D maps of WMH. Techniques of this second type are exemplified by (Price, Schmalfuss, & Sistrom, 2005b) and (Achiron, Gicquel, Miron, & Faibel, 2002), in which a rater uses image visualization software such as Scion Image (Scion Corporation, Frederick, MD) to view each slice of an MRI, and on each slice interactively adjusts a minimum and maximum threshold to visually highlight a continuous range of pixel intensities that the rater judges to represent WMH. The quantity of the slice's highlighted pixels is then summed by Scion Image, and the user can subsequently obtain a per-slice volume of WMH by multiplying the number of

highlighted pixels by the physical tissue volume (voxel volume) represented by each pixel. Per-brain WMH volumes are then obtained by summing per-slice WMH volumes. Though this does provide a per-brain WMH volume, there are a number of methodological limitations, including the inability to save the resulting WMH rating as a 3D mask, leaving raters unable to implement spatial measures of reliability, or integrate results with other spatial tissue maps such as individual or atlas-based anatomical regions. Additionally, this process can be error-prone when performed in software packages such as Scion Image, due to confusion about change in pixel counts when a rater zooms in or out of an image, or application of incorrect voxel volume when multiplying pixel count to produce a physical tissue volume. Furthermore, Scion Image lacks a robust macro language, requiring raters to perform a number of time-consuming and error-prone digital housekeeping tasks that are automated in the method developed for the current work (file naming, image conversions, volume calculations, summing across slices, etc).

**Merit: allows integration with other spatial maps**

Creating 3D WMH maps as part of a severity rating protocol confers the ability to examine endorsed WMH lesions in space together with other data that also exist in the MRI space. Examples include T1 intensity data and masks of gray matter structures, T2 FLAIR intensity data of WMH and surrounding WM, and diffusion-weighted image (DWI)-derived data such as white matter tract masks and fractional anisotropy values. These combinations facilitate both the anatomical localization of WMH lesions and further investigation of the tissue properties of the WMH lesions.

White matter contains the axons projecting from neuron cell bodies to their targets, and some of those connections are more fragile than others, with specific effects when

lesioned. Identifying *which* of an individual's white matter tracts are affected by WMH lesions is made possible by the 3D format of the maps output by segmentation: the resulting 3D map of an individual's WMH tissue can be aligned with either a) a priori standard templates of white matter tracts (Mori et al., 2008; Mori & van Zijl, 2007), or b) white matter tracts derived from the individual's own DWI data (Mahon et al., 2009; Nguyen et al., 2005). This allows analysis of the relationship between cognition and the identity of affected white matter tracts with specificity and reliability that is impossible with subjective visual rating scales.

When this WMH localization is performed through alignment of an individual's WMH lesion map with their *own* DWI data, an additional set of analyses becomes possible: fractional anisotropy, mean diffusivity, and other indices extracted from DWI data provide additional information about the health of the white matter inside and outside of the WMH regions, augmenting the T2-FLAIR-derived WMH maps. Similarly, alignment of the original T2 FLAIR intensity data with the WMH maps allows comparison of the T2 FLAIR intensity distributions of lesioned and non-lesioned white matter, as well as the relationship of those intensity distributions to the lesioned and non-lesioned intensity distributions of DWI-derived indices such as fractional anisotropy and mean diffusivity. Alignment can be achieved through careful acquisition in which participant movement between scans is minimized, or can be performed post-acquisition using manual alignment or automated computer alignment algorithms.

Three-dimensional mapping of WMH also facilitates localization of WMH lesions with respect to T1-derived brain structures, such as subcortical gray matter nuclei, specific gyri and sulci, divisions of the ventricular system, and lacunar infarcts. As WMH

represents insult to the brain's communication system, proximity to functional structures visualized on T1 images may help explain selective deficits, and proximity to other pathological markers (lacunar infarcts, enlarged ventricles, wide sulci) may aid understanding of relationships among pathological processes. Alignment to whole-brain T1 anatomical images also allows for automated analysis of WMH distribution among gross regional divisions of the brain derived from T1 images such as frontal/temporal/parietal/occipital lobe divisions, periventricular/deep/juxtacortical divisions, and anterior/posterior or dorsal/ventral divisions.

**Merit: resulting maps can be used to train people and computers**

Because the product of 3D WMH lesion mapping is a 3D map that is easily visualized as a overlay on T2 FLAIR images from which the map was derived, the data implicitly generated by every rating also serve as ready visual training material for new raters. An expert rater's every judgment about tissue boundaries or intensity differences is captured in the resulting WMH map, allowing for careful inspection and replication by trainee raters. This stands in stark contrast to the time-consuming co-rating usually necessary when training a new rater to reliability on subjective visual rating scales. Provided with only a brain's expert Junque rating and the original T2 FLAIR image, a rater-in-training is unlikely to learn much about the specific decisions made during the rating without a detailed explanation from an expert rater. In addition to providing training materials for new raters, the maps produced by new raters themselves also serve as easy-to-audit data that can be inspected and corrected by a supervising rater.

The 3D maps produced during expert ratings can also be combined with their original T2 FLAIR intensity data to train machine-learning algorithms in the performance of automated WMH segmentation (Lao et al., 2008). This approach combines the

advantages of human expertise with the efficiency of full computer algorithm automation. This training is one approach to an evolving set of computer techniques that generate automated or semi-automated WMH masks using a variety of algorithms.

**Automated Lesion Mapping: an Evolving Alternative**

For at least the last ten years, diverse groups of investigators have been developing computer algorithms designed to automatically segment WMH tissue from normal-appearing brain tissue (Achiron et al., 2002; Admiraal-Behloul et al., 2005; Anbeek, Vincken, van Osch, Bisschops, & der Grond, 2004; Beare et al., 2009; de Boer et al., 2009; Dyrby et al., 2008; Gibson, Gao, Black, & Lobaugh, 2010; Herskovits, Itoh, & Melhem, 2001; Jack et al., 2001; Jongen et al., 2009; Lao et al., 2008; Maillard et al., 2009; Sachdev, Parslow, Wen, Anstey, & Easteal, 2009; Sachdev, Wen, Chen, & Brodaty, 2007; Wen & Sachdev, 2004a; Wu et al., 2006). This is motivated by the need to generate the rich spatial information contained in 3D WMH maps, and to obviate the relatively time-consuming process of manual WMH tracing normally required to create those maps. The reliable, efficient WMH maps manually generated by the 3D mapping technique described in the present work can aid the development of these automated methods in at least two ways. First: manually created 3D binary maps of WMHs can be used as training input for automated machine-learning algorithms. Some of the automated approaches under development require the input of such training models in order to learn how expert human raters separate WMH from the surrounding tissue. Second, the 3D WMH maps manually produced with the technique presented here can be used as gold-standard human expert data against which researchers can evaluate the results of their automated segmentations.

## Study Rationale

Though severity of WMH is an important diagnostic criterion in vascular dementia, it is typically characterized using only semi-quantitative subjective visual rating scales, which have a number of significant limitations including poorly characterized reliability and low-fidelity output that communicates little detail about expert raters' judgments. In the present work the author aims to 1) develop a novel, efficient implementation of three-dimensional (3D) manual WMH mapping with high inter- and intrarater reliability, and 2) evaluate the convergent validity of this novel technique and a widely-used WMH visual rating scale. It is essential that clinicians using imaging in assessment and diagnosis choose imaging methods with high, well-characterized reliability and validity. This is also true for investigators making assertions about the relationship between biology and cognition, who bear the additional burden of extracting as much detail as possible from their imaging data to explain variance in clinical and cognitive variables.

## Aims and Hypotheses

In the interest of advancing the methods used to investigate the relationship between WMH and cognitive decline, the present work is organized around two aims:

- AIM 1: Develop a novel, efficient three-dimensional WMH manual mapping technique with high inter- and intrarater reliability. HYPOTHESIS 1: High reliability will be demonstrated within and between two raters, as measured by intraclass correlations (ICCs) of tissue volumes, and, more importantly, a measure of agreement based on spatial overlap of endorsed WMH.
- AIM 2: Evaluate the convergent validity of the novel manual 3D WMH mapping technique and an established visual rating scale. HYPOTHESIS 2: Spearman's rho will reflect acceptable reliability between the 3D WMH mapping and an established visual rating scale.

If reliable, the authors propose that 3D digital segmentation of WMH should supplant subjective visual rating scales when possible.

The manual 3D digital segmentation technique developed for the present work was built upon the image processing and display capabilities of the NIH software package ImageJ and its macro language (Rasband, 1997). This allowed the creation of an interface that guides raters through WMH segmentation steps efficiently, while minimizing opportunities for raters to omit essential steps, misname output files, or accidentally overwrite existing images. Details are provided in the methods section below.

# CHAPTER 2
# METHODS

In a cross-sectional observational study of repeated measures, two expert raters used a novel technique to manually create 3D digital maps of WMH tissue from 15 whole-brain MRIs previously scored with the Junque subjective visual rating scale (Libon et al., 2008). Inter- and intrarater reliability of these 3D segmentations were evaluated with 1) a measure of spatial overlap (Dice similarity coefficient) and 2) intraclass correlation (ICC) of the WMH tissue volumes computed from manual 3D segmentations. Convergent validity with previously acquired Junque subjective visual ratings was evaluated by Spearman's rank correlation coefficient. To further evaluate the convergent validity of methods, one rater (SM) then created manual 3D digital maps for the remaining cases in the original dataset of 86, 77 of which had been previously assigned Junque subjective visual rating scores. This allowed convergent validity of methods to be evaluated across a larger dataset (n=77) than the 15 cases originally chosen for inter- and intrarater reliability of WMH volumes.

## Participants – Brain Magnetic Resonance Images

The present work is a secondary analysis of single-time-point brain MRIs collected from dementia patients at the University of Medicine and Dentistry in New Jersey (UMDNJ). Informed written consent was obtained according to UMDNJ and University of Florida Institutional Review Board guidelines and the Declaration of Helsinki. For the present repeated-measure reliability analysis, 15 participants' T2 FLAIR brain MRIs were chosen from a larger set of 86 two-dimensional (2D) T2 FLAIR brain MRIs (1.5T Siemens MRI: 26 slices, 5 mm thick, 1 mm gap), which were previously acquired from research participants diagnosed with dementia and enrolled in an ongoing investigation

of WMH severity and cognition (Libon et al., 2008). This larger set of 86 brain images were drawn from a sample of > 100 consecutive intakes at the UMDNJ Memory Disorders Clinic who were free of stroke and met criteria for probable AD per National Institute of Neurological Disorders and Stroke, Association Internationale pour la Recherché et l'Enseignement en Neurosciences (NINDS-AIREN). Demographic data are summarized in Table 2-1. The 15 images analyzed in the present work were selected based on their previously assigned Junque WMH severity ratings: five of each mild, moderate, and severe WMH cases were selected in order to maximize the range of WMH severity examined. Each of these 15 original MRIs were duplicated and pseudorandomly reblinded, so that each rater measured each of the 15 MRIs in two non-consecutive ratings while blind to the identity of repeated ratings and all previously acquired data. This subset was limited to 15 cases as a practical limitation of rater availability: (15 cases) X (two ratings per case) X (projected 15 minutes per rating) = 7.5 hours of projected data collection per rater.

Table 2-1. Demographics of brain magnetic resonance image participants

| MRI dataset and subsets | *n* | Female | Age in years *M* (*SD*) | Education in years *M* (*SD*) |
|---|---|---|---|---|
| Available cases, WMH-mapped by 1 or 2 raters | 86 | 65 | 80.06 (5.45) | 12.05 (2.89) |
| Subset with previously assigned Junque scores | 77 | 55 | 80.25 (5.49) | 11.99 (3.00) |
| Subset mapped by 2 raters twice for reliability | 15 | 11 | 82.76 (3.64) | 12.60 (3.29) |

*Note*. *M* = mean; MRI = magnetic resonance image; *n* = number of cases; *SD* = Standard Deviation; WMH = white matter hyperintensity.

## Existing Subjective Ratings of White Matter Hyperintensities

During data collection for previous work (Libon et al., 2008),each brain scan received a subjective visual rating of WMH severity by non-radiologist expert raters trained to 99% reliability on the Junque subjective visual rating scale for WMH (Junque et al., 1990). The method, resulting in a semi-quantitative rating of 0 to 40 per brain, is described on page 152 of the Junque 1990 article:

> A method for the systematic quantification of LA in MRI was developed using a modification of Rezek and coworkers' criteria. The following five areas from the left and right sides were assessed: centra semiovale on the frontal region, centra semiovale on the parietal region, white matter surrounding the frontal horn of the lateral ventricle, white matter surrounding the corpus of the lateral ventricle, and white matter surrounding the atrium and occipital horn of the lateral ventricle. This analysis included the white matter regions where LA is common. Each area was evaluated on the coronal and transverse slices and the extension of white matter that appeared hyperintense on the T2-weighted images was assessed. Numerical scores of 0 through 4 were assigned for each area: 0, no changes; 1, <25% hyperintense; 2, 25% to 50% hyperintense; 3, 50% to 75% hyperintense; and 4, >75% hyperintense. An LA score of 0 to 40 was possible for each patient (5 brain areas X 2 hemispheres, with a maximum score of 4 per area).

Though there are a number of other WMH subjective visual rating scales in the literature, the Junque scale confers analytic flexibility that others do not: Junque scores are interval data that do not suffer from restricted range (Cosentino, Jefferson, M. Carey, Price, Davis Garrett, Swenson, et al., 2004a), and have psychometric properties shown to be appropriate for both dependent and explanatory variables (Libon et al., 1998).

## Novel Manual Segmentation of White Matter Hyperintensities

Raters CP and SM, who were proficient in Junque subjective visual ratings of WMH and digital brain tracing tools, were trained to produce 3D segmentations of WMH tissue using the novel manual segmentation method described below. They each then

applied this method to create 3D WMH segmentations of the 15 randomized and re-blinded brain volumes in the sample, for 30 ratings per rater (15 cases x two ratings). One rater (SM) then applied this measurement technique to the remaining cases of the 77 that had been assigned Junque subjective visual rating scores in previous work.

Using ImageJ (Rasband, 1997) on Linux workstations, raters identified and "segmented" WMHs to produce a 3D digital tissue map, or "binary mask" (Figure 2-1). (The term "binary" here refers to the way WMH endorsement is coded in the resulting map: all points in MRI space are represented as black background with intensity value zero, except for areas endorsed as WMH, which have an intensity value one.) Locally-developed ImageJ macros reduced manual segmentation to less than15 minutes per FLAIR volume, performed in three steps per axial slice: 1) adjust brightness and contrast for subjectively best contrast between WMH and normal-appearing tissue, 2) use a computer mouse to draw gross, liberal boundaries around the regions containing WMHs, 3) highlight only the hyperintense voxels by adjusting upper- and lower-threshold intensity sliders. Voxels that are both inscribed and highlighted are assigned intensity value one in the slice's 2D WMH binary mask. This was repeated for each axial slice in the brain, and the resulting 2D per-slice masks were concatenated to form a single 3D binary mask for each brain.

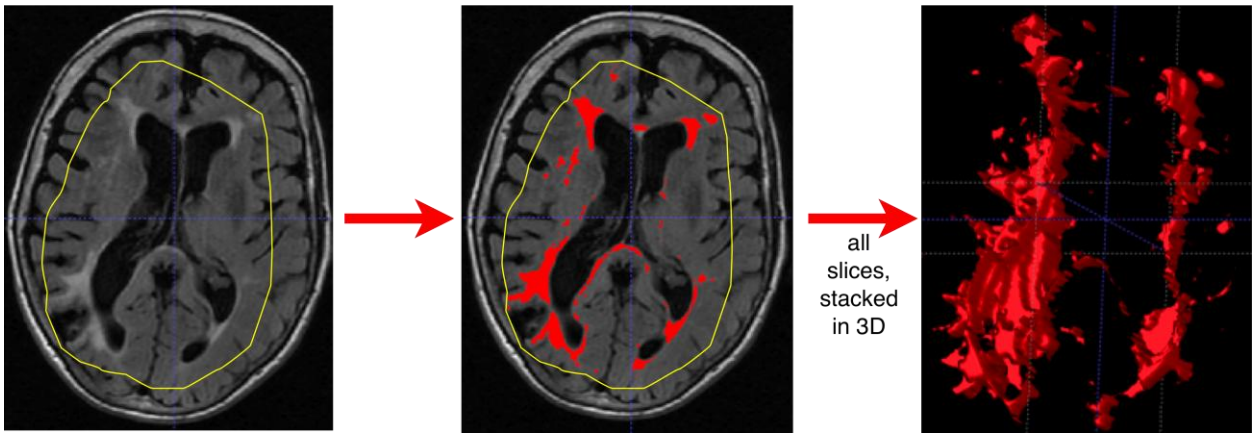


Figure 2-1. Raters manually segmented white matter hyperintensities from normal-appearing tissue on individual two-dimensional slices, which were then concatenated into a three-dimensional binary mask for each rating.

## Statistical Evaluations of Reliability

### Inter- and Intrarater Reliability of White Matter Hyperintensity Segmentation

Inter- and intrarater reliability of manual WMH segmentations were evaluated by calculating an index of spatial overlap, the Dice similarity coefficient (DSC), for pairs of repeated ratings. It is calculated as the overlap volume common to two ratings divided by one half the sum of their volumes. A DSC value of 1 represents perfect overlap, 0 represents no overlap, and values between 1 and 0 are interpreted like kappa, with > .7 indicating excellent agreement (Zijdenbos et al., 1994). This intuitive interpretation is why DSC was chosen from among a number of other reliability metrics that can be applied across 3D space. All calculations were performed by a locally written script incorporating FSL command-line programs fslmaths and fslstats (S. M. Smith et al., 2004; Woolrich et al., 2009).

A second measure of inter- and intrarater reliability, intraclass correlation (ICC), was calculated for the WMH mask *volumes*, providing a more commonly used reliability coefficient with which to compare DSC. Intraclass correlation was calculated as a two-

way random effects model for absolute agreement using SPSS 17 for MS Windows (McGraw & Wong, 1996; Shrout & Fleiss, 1979). The ICC analysis does not account for degree of spatial overlap among repeated ratings.

**Convergent Validity of Methods**

Finally, the convergent validity of Junque subjective visual rating and 3D manual segmentation of WMH was evaluated by using Spearman's rank correlation coefficient to characterize the reliability between Junque subjective visual ratings and WMH volumes calculated from 3D manual segmentation. Spearman's rho was chosen due to the great difference in the methods' range and psychometric properties, which renders parametric statistical comparisons inappropriate. A scatter plot of ranks reflects this nonparametric analysis..

# CHAPTER 3
# RESULTS

High reliability was observed within and across raters (aim 1), and between the two WMH quantification techniques (aim 2), as detailed below.

## Aim 1: Inter- and Intrarater Reliability of White Matter Hyperintensity Segmentation

High inter- and intrarater reliability of manual WMH segmentations were observed: large DSCs indicated high proportions of spatial overlap in repeated ratings (Table 3-1). The mean DSC for all 90 inter- and intrarater comparisons was .837 (*SD* = .115, minimum =. 39, maximum = .98.). High inter- and intrarater agreement was also observed in WMH volumes calculated from WMH maps: ICC was .960 across all four sets of ratings (95% Confidence Interval .882 to .986; Cronbach's Alpha = .994).

Table 3-1. Manual three-dimensional digital white matter hyperintensity segmentations are reliable within and between raters.

| | Rater 1, rating a DSC: *M* [95% CI] | Rater 1, rating b DSC: *M* [95% CI] | Rater 2, rating a DSC: *M* [95% CI] | Rater 2, rating b DSC: *M* [95% CI] |
|---|---|---|---|---|
| Rater 1, rating a | 1 | -- | -- | -- |
| Rater 1, rating b | .933 [.917, .950] | 1 | -- | -- |
| Rater 2, rating a | .831 [.800, .862] | .827 [.770, .884] | 1 | -- |
| Rater 2, rating b | .798 [.730, .866] | .792 [.710, .873] | .842 [.765, .918] | 1 |

Note. CI = confidence interval; DSC = Dice similarity coefficient; *M* = mean. Ratings a and b refer to a rater's repeated ratings of the 15 MRIs. Cells contain mean DSC and 95% confidence interval across all 15 MRIs. DSC >.7 indicates highly reliable repeated ratings per spatial overlap.

## Aim 2: Convergent Validity of Methods

High convergent validity was observed between the previously acquired Junque subjective visual rating scores and the WMH volumes for all sets of WMH masks (Spearman's rho in Table 3-2). A scatter plot depicts the bivariate relationship of these measurements for all 77 cases (Figure 3-1), as does a scatter plot of their ranks (Figure 3-2), which more closely represents the relationship expressed by Spearman's rho. Descriptive statistics for both measures are summarized in Table 3-3.

Table 3-2. White matter hyperintensity map volumes are reliable with previously assigned Junque scores.

| Data subset | *n* | WMH map volume correlation with previously acquired Junque scores (Spearman's rho) |
|---|---|---|
| Rater 1, rating a | 15 | .845 ($p < .01$) |
| Rater 1, rating b | 15 | .858 ($p < .01$) |
| Rater 2, rating a | 15 | .849 ($p < .01$) |
| Rater 2, rating b | 15 | .831 ($p < .01$) |
| Rater 1, ratings of all cases that had previously assigned Junque scores | 77 | .760 ($p < .01$) |

Note. *n* = number of cases; WMH = white matter hyperintensity. Ratings a and b refer to a rater's repeated ratings of an image set.

Table 3-3. Summary statistics of Junque scores and white matter hyperintensity map volumes analyzed for convergent reliability of methods (n = 77)

| Measure | *n* | *M* (*SD*) | Minimum | Maximum |
|---|---|---|---|---|
| Junque subjective visual rating scale (score) | 77 | 9.87 (7.25) | 0 | 37 |
| WMH map volume (microliters) | 77 | 20818.08 (18581.63) | 1833.40 | 117268.30 |

Note. *M* = mean; *n* = number of cases; *SD* = standard deviation; WMH = white matter hyperintensity.

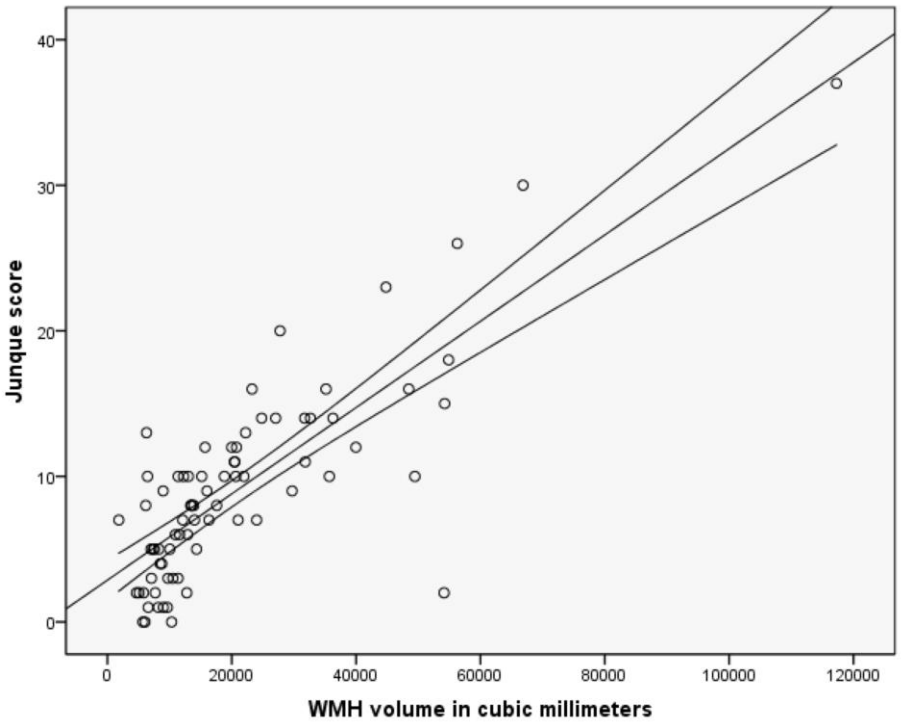


Figure 3-1. Scatter plot, best-fit line, and 95% confidence interval of Rater 1 white matter hyperintensity volumes and previously acquired Junque scores (n=77).

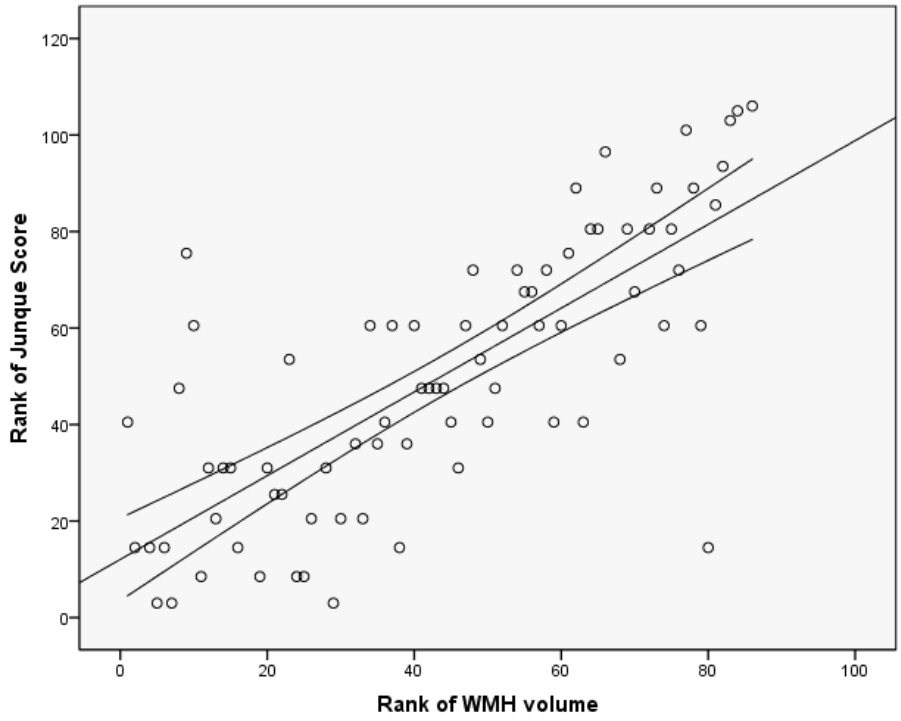


Figure 3-2. Scatter plot, best-fit line, and 95% confidence interval of Rater 1 white matter hyperintensity volume ranks and ranks of previously acquired Junque scores (n=77).

# CHAPTER 4
# DISCUSSION

Using subjective visual rating scales, investigators have previously established that WMH severity is a clinically and cognitively significant neuroimaging sign (Libon et al., 2008; Silbert et al., 2009). By demonstrating convergent validity with the well-established Junque subjective visual rating scale, the present study provides support for conducting future T2-FLAIR-based WMH measurements with manual 3D digital segmentation, a precise tissue measurement technique with a number of methodological advantages over subjective visual rating scales. This technique, demonstrated to be reliable and efficient in the present work's novel implementation, produces a perfect 3D digital map of the tissue a rater endorses as WMH. This high-fidelity output confers a number of advantages over subjective visual rating scores. First is the detailed characterization of repeated rating reliability by analysis of the mutual spatial overlap of the 3D WMH maps produced. This not only includes the intuitive visual inspection of areas of non/overlap in repeated ratings, but also the *quantification* of repeated-measure reliability by an index of spatial overlap such as the Dice similarity coefficient. A second advantage over subjective visual rating scales is the ability to calculate a physical WMH tissue volume from each 3D WMH map. This allows WMH severity to be expressed in proportion to other physical tissue volumes such as total white matter volume or total intracranial volume, and produces data free of psychometric problems identified in some subjective visual rating systems. A third advantage over subjective visual rating scales is the ability to combine the resulting 3D WMH map with other 3D brain data that exist in MRI space. This facilitates both the anatomical localization of WMH tissue (e.g., by three-dimensionally superimposing

individual anatomical maps or standard atlases), and deeper investigation of WMH tissue properties (e.g., by examining the MR-diffusion properties of expert-endorsed WMH tissue). A fourth advantage over subjective visual rating scales is that 3D WMH maps are easily implemented in training and auditing new raters and machine-learning algorithms. This obviates the intensive co-rating and full-time supervision usually necessary to train a new rater to reliability with a subjective visual rating scale: every expert manual 3D WMH rating produces output that serves as a readily-inspected 3D gold-standard model against which a trainee can compare their own ratings. Similarly, 3D WMH maps from expert manual ratings can be used as input to train WMH segmentation machine-learning algorithms, and as a gold-standard against which the results of automated computer algorithm WMH mapping techniques can be compared.

## Clinical Relevance of White Matter Hyperintensity Severity

The present work is motivated by a desire to improve tools for investigating the relationship between damage to the brain's circuitry and individuals' abilities and experiences. The last forty years of clinical research has revealed that WMH is both a common incidental finding in normally functioning individuals and a biomarker linked with clinical and cognitive decline in patients (Libon et al., 2004). Improving WMH quantification techniques may help describe the complicated relationships among clinical, cognitive, demographic, and physiological variables.

### White Matter Hyperintensity is Associated with Clinical and Cognitive Deficits

Though healthy older adults commonly present with WMH as an incidental finding, older adults with cognitive impairment present with WMH more frequently and with more severe WMH. This is especially true when cardiovascular risk factors are comorbid with their cognitive symptoms (Junque et al., 1990). The received science so convincingly

links WMH and cognitive decline that presence and severity of WMH in a cognitively compromised patient is a criterion for the diagnosis of Vascular Dementia (VaD) under many classification schemes (Cosentino, Jefferson, M. Carey, Price, Davis Garrett, Swenson, et al., 2004a). Existing literature suggests that location, severity, and progression of WMH severity may explain significant portions of the variance in WMH-linked cognitive deficits. Our hope is that these relationships may be further clarified by the precise WMH measurements obtained from the technique presented in the current work.

**Severity matters**

The high prevalence of WMH as an incidental finding in asymptomatic aging adults suggests that the mere presence of WMH is not a sensitive or specific predictor of cognitive decline. Previous works suggest that there may be a threshold effect in which some proportion of the brain's white matter must be affected before cognition is impacted (Erkinjuntti et al., 1987; Erkinjuntti, Haltia, Palo, Sulkava, & Paetau, 1988). In fact, NINDS-AIREN research criteria for VaD require that at least "one quarter" of a brain's white matter be affected by WMH for the patient to be classified as VaD (Román et al., 1993). Unfortunately these severity thresholds are based on subjective visual rating scale data, which are subject to the limitations described above. The manual 3D WMH mapping method presented in the current work can be used to confirm or disconfirm hypothesized thresholds by providing a valid and reliable measure of WMH tissue volume, which can also be easily expressed as a physical proportion of total white matter volume, total parenchyma volume, or total intracranial volume.

Severity of WMH may also have a relationship with the *type* of deficits experienced by a patient. Price et al. (2005) examined the relationship between dementia symptom

profile and WMH severity. In a sample of probable AD patients the authors confirmed that the patient group with the most severe WMH had primarily executive-visuoconstruction deficits, while the group with the least severe WMH had primarily memory and language deficits. (The group with only moderate WMH fell somewhere in between, with executive deficits only slightly worse than memory-language deficits). This association between more severe WMH and more severe executive dysfunction is also supported by an analysis of clock-drawing errors: in a comparison between low-WMH individuals and high-WMH individuals, the high-WMH individuals produced more clock executive errors (Time, Spatial Layout, Preservation/Pull to Stimulus subscales) than the low-WMH group (Cosentino, Jefferson, Chute, Kaplan, & Libon, 2004b). Unfortunately these findings were derived from WMH subjective visual rating scale data, the distributions of which necessitated categorical analysis of high- and low-WMH groups instead of parametric analysis of WMH as a continuous variable. The manual 3D WMH mapping technique presented in the current work can be used to provide volumetric measures of WMH, avoiding these distributional limitations and facilitating parametric analyses of the relationship between WMH severity and specific cognitive deficit profiles.

The predictive value of volumetric measurement of WMH tissue is supported by a recent longitudinal study of 307 community-dwelling elderly individuals (Carmichael et al., 2010). Measures of WMH volume, episodic memory, semantic memory, and executive function were analyzed. Individuals with less severe WMH at baseline demonstrated slower declines in executive function and semantic memory over the subsequent four years. This work is of particular value due to its longitudinal design and

the construction of its large and diverse sample: the WMH volume finding survives a number of demographic variables that are frequent confounds in smaller cross-sectional studies, including ethnicity, socioeconomic status, education, baseline cognitive function, cardiovascular health, early life experiences, and environmental exposures. Though these results are based on WMH volumes generated by automated (histogram-based) computer algorithm segmentation of WMH from non-WMH white matter, the manual 3D WMH mapping technique presented in the current work can be used to replicate and extend this important volume-based analysis with WMH volumes produced by expert manual mappings of WMH.

**Location matters**

Given the presumed vascular etiology of WMH, it is not surprising that the distribution of WMH across the brain varies just as vascular supply varies throughout the brain. In the literature first recognizing WMH as a neuroimaging sign, authors noted that it tended to be more present in periventricular white matter than other white matter regions (De Reuck, 1971). Modern neuroimaging techniques now permit a higher-resolution analysis of the topographical distribution of WMH, and thus far the findings are largely consistent with the historical view that WMH appears most commonly in the periventricular white matter surrounding the lateral ventricles and deep subcortical nuclei (e.g., thalamus, caudate), but is also detectable in the juxtacortical (u-fiber) regions of white matter (Geurts et al., 2005). As neuroimaging technology improves we are also able to more closely examine the overlap in vascular and WMH territories: in one recent high-resolution MRI study of 477 healthy participants aged 60-64 y, authors found that the greatest concentration of WMH was in the territories of the lenticulostriate arteries (Wen & Sachdev, 2004b).

It is reasonable to expect that the cognitive effects of WMH vary as a function of location in the brain, just as the role of white matter varies with location in the brain. The juxtacortical white matter is largely composed of short mu-fibers that connect areas of local cortex, while deep periventricular white matter is largely composed of long cortical-subcortical, subcortical-spinal, cortical-spinal, and interhemispheric connections. It is reasonable to hypothesize that cognitive domains associated with cortical gray matter, such as language and immediate and delayed memory, would be not be as affected by periventricular and deep WMH as they would juxtacortical WMH that interferes with cortical association networks. Similarly, it is unlikely that domains linked with the subcortical nuclei, such as language, motor performance, processing speed, and executive function, would be affected by juxtacortical WMH (Alexander, DeLong, & Strick, 1986). Recent work supports these hypotheses, with one analysis of the relationship between regional WMH and a number of cognitive domains finding a significant relationship between measures of attention and processing speed, and the amount of WMH surrounding the ventricles and deep cortical nuclei (Tate et al., 2008). Another recent study found that higher-level working memory tasks are more impacted by left-hemisphere WMH than right-hemisphere WMH, and that even within the left hemisphere, higher-level working memory mental manipulations of disengagement and temporal re-ordering were disrupted by WMH in the frontal centrum semiovale but not by more anterior WMH surrounding the frontal horn of the left ventricle (Lamar et al., 2008).

Careful analysis of factors related to the *location* of WMH tissue requires accurate anatomical localization of WMH tissue. The manual 3D WMH mapping technique

presented in the current work provides a 3D map of the tissue a rater endorses as WMH. Because this map exists in MRI space it is easily combined with localizing data that also exists in MRI space, including vascular territory maps generated by MR angiograms, white-matter tract maps generated from an individual's MR-diffusion data, and standard atlas-based maps of gray and white matter regions. This provides researchers with a number of objective, reliable approaches for characterizing the anatomical location of WMH tissue.

**Progression matters**

Refining WMH measurement methods is especially important given recent findings that the *progression* of WMH is a better predictor of cognitive impairment than *baseline* WMH (Silbert, Nelson, Howieson, Moore, & Kaye, 2008). Further supporting the importance of accurate longitudinal quantification of WMH is the possibility that WMH tissue volume may be a useful biomarker in prospective risk assessment: a 2009 study found that every 1 mL per year increase in periventricular WMH volume was associated with a 94% increased risk of persistent cognitive impairment (Silbert et al., 2009). Reliable WMH measurement techniques are required for replication of these sensitive longitudinal analyses, and unfortunately recent findings suggest that that subjective visual rating scales of WMH are not sufficiently reliable or sensitive for measuring these longitudinal WMH changes (den Heuvel et al., 2006; Gouw et al., 2008; Kapeller et al., 2003; Prins et al., 2004). In contrast, the preset work demonstrates that manual 3D WMH mapping provides a reliable measure of WMH tissue volume similar to the WMH volumes generated for these important longitudinal findings.

## White Matter Hyperintenisty is Associated with Advancing Age and Cardiovascular Risk Factors

Though WMH severity is currently considered a biomarker associated with vascular dementia, WMH severity is not exclusively a marker of cognitive decline. WMH has independent associations with age and cardiovascular health. Researchers have long recognized that the WMH observed on CT and MRI is frequently incidental and can be observed without neurological or cognitive sequelae (I. A. Awad, Spetzler, Hodak, C. A. Awad, & R. Carey, 1986). A 1995 literature review (Pantoni & Garcia, 1995) revealed almost unanimous agreement that frequency and severity of WMH increase as adults age, and that the increase is independent of cardiovascular risk factors in persons older than 60 years. Depending on the population being studied, authors have found that WMH is present in 15% to 65% of adults (Breteler et al., 1994; Liao et al., 1996; Lindgren et al., 1994; R. Schmidt et al., 1993; R. Schmidt, Fazekas, Kapeller, H. Schmidt, & Hartung, 1999; Tomimoto et al., 1996; Ylikoski et al., 1993), with a three-fold increase In older relative to younger adults (Hogervorst, Ribeiro, Molyneux, Budge, & A. D. Smith, 2002). The relationships among cognitive decline, WMH, aging, and cardiovascular risk factors could be explained by a single three-component model in which 1) cognitive function is negatively impacted by cumulative white matter damage, which can be 2) caused by cerebrovascular pathology, which 3) increases with both advancing age and cardiovascular risk factors.

This model is thus far supported by post-mortem histopathological studies of WMH, which have led to wide acceptance that WMH is a marker of white matter damage caused by disruptions in vascular integrity. A number of cellular pathologies have been observed in WMH, including demyelination, axon loss, gliosis, and

spongiosis (Chimowitz, Estes, Furlan, & I. A. Awad, 1992; Munoz, Hastak, Harper, Lee, & Hachinski, 1993; Scarpelli et al., 1994). In a 2008 histopathological study of WMH in 20 consecutive cases from the Australian Brian Donor programs, Young et al. examined areas of WMH for histopathological signs of damage and for vascular integrity (measured by the presence of CD31, a major component of the cell junctions that form the blood-brain barrier). They found that decreased vascular integrity was the best predictor of increased WMH severity, and that WMH tissue had abnormally low expression of P-gp, a molecular pump that maintains the internal environment of the brain by transporting unwelcome substances from the brain back into the bloodstream. This histopathological evidence supports the theory that increased permeability of intracranial vasculature is the final common pathway by which a number of cardiovascular risk factors cause damage to the brain's white matter. Diabetes mellitus, smoking, hypertension, hypercholesterolemia, obesity, atherosclerosis, and embolic cerebral injury can all impact the functioning of the blood-brain barrier in this way, thereby creating areas of damaged white matter that appear as WMH. Replicating and extending these histopathological findings requires a precise 3D map of WMH tissue in order to prescribe the boundaries for histopathological analysis. The method presented in the current work is an efficient, reliable technique with which an expert rater may create 3D maps of WMH tissue to be used in postmortem histopathological analysis of the processes underpinning the development of WMH.

## Ongoing Extension

The present work is currently being extended in a number of ways. First, the relationship between cognitive decline and the anatomical location of WMH is being investigated. Automated methods separate each brain's white matter into three mutually

exclusive compartments: periventricular, deep, and juxtacortical. By superimposing these regions on the 3D WMH maps produced in the present work, we are able to examine the relationship between cognitive performance and the regional volume (i.e., location) of WMH. This general approach of intersecting WMH maps with a priori regions of interest is equally applicable to analyzing the intersection of WMH maps and anatomical regions derived from any combination of MR imaging (DWI, FMRI, arterial spin labeling) and brain segmentation technique. Second, we are examining interactions with gross anatomical variables, including the volume of the intracranial vault, total brain volume, and total white matter volume. It is possible that relationship between WMH and cognitive decline is most clear when WMH volume is expressed as a percentage of one of these larger volumes, rather than a raw tissue volume. It is easy to imagine how some large N cubic millimeters of WMH might impact a given brain less than it might impact a much smaller brain. This proportional expression of endorsed WMH volume is especially important for achieving parity with clinically-validated subjective visual rating scales that employ scores representing the proportion, not volume, of white matter affected.

## Limitations

A technical limitation of the technique presented here is that it is not presently capable of presenting the user with images from multiple types of MRI to aid the creation of WMH maps. Though T2 FLAIR images are the accepted standard for making clinical decisions based on WMH, white matter imaging methods based on MR Diffusion Weighted Imaging (DWI) are maturing quickly, and will likely come to augment T2 FLAIR images in clinical judgments of white matter health. Once a DWI-derived metric, such as fractional anisotropy or mean diffusivity, is found to be useful in WMH

analysis, the method presented here can be extended to include those additional images either by presenting DWI-derived images adjacent to the FLAIR images, or by overlaying the two.

The present work was conducted under an untested assumption common to the visual rating scale and pixel-counting work that preceded it: that the single-time-point T2 FLAIR images analyzed are a reliable and stable indicator of actual tissue condition at the time of image acquisition. Water content affects the MR signal used to create FLAIR images, and it is possible that water concentration and other physiological variables may produce variability in FLAIR results that confounds the quantitative analysis of cross-sectional single-time-point images. Serially acquired scans repeated within days or weeks of each other would confirm or disconfirm this assumption, but they were not part of the original protocol from which these scans were acquired.

The present work was performed on relatively low resolution T2 FLAIR images (26 slices per brain). Though this resolution is commonly acquired for clinical and research studies, newer high-resolution T2 FLAIR sequences with gapless slices as thin as 1 mm will become faster and easier to acquire as MRI technology progresses. At 5 to 30 seconds per slice, the manual 3D WMH mapping method presented here will require substantially more time for a 160-slice high-resolution T2 FLAIR than the 26-slice images measured in the present work.

Though the methods presented in the present work may be equally useful in research and clinical settings, clinical use within the United States would first require FDA clearance as a Class II medical device. An efficient solution would be the re-

implementation of the present methods using existing FDA-cleared software. Interested investigators should seek advice from their institution's regulatory specialists.

LIST OF REFERENCES

Achiron, A., Gicquel, S., Miron, S., & Faibel, M. (2002). Brain MRI lesion load quantification in multiple sclerosis: a comparison between automated multispectral and semi-automated thresholding computer-assisted techniques. *Magnetic resonance imaging*, *20*(10), 713–720.

Admiraal-Behloul, F., den Heuvel, van, D. M. J., Olofsen, H., van Osch, M. J. P., der Grond, van, J., van Buchem, M. A., & Reiber, J. H. C. (2005). Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. *NeuroImage*, *28*(3), 607–617. doi:10.1016/j.neuroimage.2005.06.061

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual review of neuroscience*, *9*, 357–381. doi:10.1146/annurev.ne.09.030186.002041

Anbeek, P., Vincken, K. L., van Osch, M. J. P., Bisschops, R. H. C., & der Grond, van, J. (2004). Automatic segmentation of different-sized white matter lesions by voxel probability estimation. *Medical Image Analysis*, *8*(3), 205–215. doi:10.1016/j.media.2004.06.019

Awad, I. A., Spetzler, R. F., Hodak, J. A., Awad, C. A., & Carey, R. (1986). Incidental subcortical lesions identified on magnetic resonance imaging in the elderly. I. Correlation with age and cerebrovascular risk factors. *Stroke; a journal of cerebral circulation*, *17*(6), 1084–1089.

Beare, R., Srikanth, V., Chen, J., Phan, T. G., Stapleton, J., Lipshut, R., & Reutens, D. (2009). Development and validation of morphological segmentation of age-related cerebral white matter hyperintensities. *NeuroImage*, *47*(1), 199–203. doi:10.1016/j.neuroimage.2009.03.055

Breteler, M. M., van Swieten, J. C., Bots, M. L., Grobbee, D. E., Claus, J. J., den Hout, van, J. H., van Harskamp, F., et al. (1994). Cerebral white matter lesions, vascular risk factors, and cognitive function in a population-based study: the Rotterdam Study. *Neurology*, *44*(7), 1246–1252.

Carmichael, O., Mungas, D., Beckett, L., Harvey, D., Tomaszewski Farias, S., Reed, B., Olichney, J., et al. (2010). MRI predictors of cognitive change in a diverse and carefully characterized elderly population. *Neurobiology of aging*. doi:10.1016/j.neurobiolaging.2010.01.021

Chimowitz, M. I., Estes, M. L., Furlan, A. J., & Awad, I. A. (1992). Further observations on the pathology of subcortical lesions identified on magnetic resonance imaging. *Archives of neurology*, *49*(7), 747–752.

Cosentino, S. A., Jefferson, A. L., Carey, M., Price, C. C., Davis Garrett, K., Swenson, R., & Libon, D. J. (2004a). The clinical diagnosis of vascular dementia: A comparison among four classification systems and a proposal for a new paradigm. *The Clinical Neuropsychologist*, *18*(1), 6–21.

Cosentino, S. A., Jefferson, A. L., Chute, D. L., Kaplan, E., & Libon, D. J. (2004b). Clock drawing errors in dementia: neuropsychological and neuroanatomical considerations. *Cognitive and behavioral neurology : official journal of the Society for Behavioral and Cognitive Neurology*, *17*(2), 74–84.

Crum, W. R., Camara, O., & Hill, D. L. G. (2006). Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE transactions on medical imaging*, *25*(11), 1451–1461. doi:10.1109/TMI.2006.880587

de Boer, R., Vrooman, H. A., der Lijn, van, F., Vernooij, M. W., Ikram, M. A., der Lugt, van, A., Breteler, M. M. B., et al. (2009). White matter lesion extension to automatic brain tissue segmentation on MRI. *NeuroImage*, *45*(4), 1151–1161. doi:10.1016/j.neuroimage.2009.01.011

De Reuck, J. (1971). The human periventricular arterial blood supply and the anatomy of cerebral infarctions. *European neurology*, *5*(6), 321–334.

den Heuvel, van, D. M. J., Dam, ten, V. H., de Craen, A. J. M., Admiraal-Behloul, F., van Es, A. C. G. M., Palm, W. M., Spilt, A., et al. (2006). Measuring longitudinal white matter changes: comparison of a visual rating scale with a volumetric measurement. *AJNR American journal of neuroradiology*, *27*(4), 875–878.

Dyrby, T. B., Rostrup, E., Baaré, W. F. C., van Straaten, E. C. W., Barkhof, F., Vrenken, H., Ropele, S., et al. (2008). Segmentation of age-related white matter changes in a clinical multi-center study. *NeuroImage*, *41*(2), 335–345. doi:10.1016/j.neuroimage.2008.02.024

Erkinjuntti, T., Ketonen, L., Sulkava, R., Sipponen, J., Vuorialho, M., & Iivanainen, M. (1987). Do white matter changes on MRI and CT differentiate vascular dementia from Alzheimer's disease *Journal of Neurology, Neurosurgery & Psychiatry*, *50*(1), 37–42.

Erkinjuntti, T., Haltia, M., Palo, J., Sulkava, R., & Paetau, A. (1988). Accuracy of the clinical diagnosis of vascular dementia: a prospective clinical and post-mortem neuropathological study. *Journal of Neurology, Neurosurgery & Psychiatry*, *51*(8), 1037–1044.

Fazekas, F., Chawluk, J. B., Alavi, A., Hurtig, H. I., & Zimmerman, R. A. (1987). MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *AJR American journal of roentgenology*, *149*(2), 351–356.

Geurts, J. J. G., Pouwels, P. J. W., Uitdehaag, B. M. J., Polman, C. H., Barkhof, F., & Castelijns, J. A. (2005). Intracortical lesions in multiple sclerosis: improved detection with 3D double inversion-recovery MR imaging. *Radiology*, *236*(1), 254–260. doi:10.1148/radiol.2361040450

Gibson, E., Gao, F., Black, S. E., & Lobaugh, N. J. (2010). Automatic segmentation of white matter hyperintensities in the elderly using FLAIR images at 3T. *Journal of magnetic resonance imaging : JMRI*, *31*(6), 1311–1322. doi:10.1002/jmri.22004

Giovannetti, T., Hopkins, M. W., Crawford, J., Bettcher, B. M., Schmidt, K. S., & Libon, D. J. (2008). Syntactic comprehension deficits are associated with MRI white matter alterations in dementia. *Journal of the International Neuropsychological Society: JINS*, *14*(4), 542–551. doi:10.1017/S1355617708080715

Giovannetti, T., Schmidt, K. S., Gallo, J. L., Sestito, N., & Libon, D. J. (2006). Everyday action in dementia: evidence for differential deficits in Alzheimer's disease versus subcortical vascular dementia. *Journal of the International Neuropsychological Society: JINS*, *12*(1), 45–53. doi:10.1017/S1355617706060012

Gouw, A. A., der Flier, Van, W. M., van Straaten, E. C. W., Pantoni, L., Bastos-Leite, A. J., Inzitari, D., Erkinjuntti, T., et al. (2008). Reliability and sensitivity of visual scales versus volumetry for evaluating white matter hyperintensity progression. *Cerebrovascular diseases (Basel, Switzerland)*, *25*(3), 247–253. doi:10.1159/000113863

Herskovits, E. H., Itoh, R., & Melhem, E. R. (2001). Accuracy for detection of simulated lesions: comparison of fluid-attenuated inversion-recovery, proton density--weighted, and T2-weighted synthetic brain MR imaging. *AJR American journal of roentgenology*, *176*(5), 1313–1318.

Hogervorst, E., Ribeiro, H. M., Molyneux, A., Budge, M., & Smith, A. D. (2002). Plasma homocysteine levels, cerebrovascular risk factors, and cerebral white matter changes (leukoaraiosis) in patients with Alzheimer disease. *Archives of neurology*, *59*(5), 787–793.

Jaccard, P. (1912). The distribution of flora in the alpine zone. *New Phytologist*, *11*, 37–50.

Jack, C. R., O'Brien, P. C., Rettman, D. W., Shiung, M. M., Xu, Y., Muthupillai, R., Manduca, A., et al. (2001). FLAIR histogram segmentation for measurement of leukoaraiosis volume. *Journal of magnetic resonance imaging : JMRI*, *14*(6), 668–676.

Jongen, C., der Grond, van, J., Anbeek, P., Viergever, M. A., Biessels, G. J., Pluim, J. P. W., & Utrecht Diabetic Encephalopathy Study Group. (2009). Construction of periventricular white matter hyperintensity maps by spatial normalization of the lateral ventricles. *Human Brain Mapping*, *30*(7), 2056–2062. doi:10.1002/hbm.20651

Junque, C., Pujol, J., Vendrell, P., Bruna, O., Jódar, M., Ribas, J. C., Viñas, J., et al. (1990). Leuko-araiosis on magnetic resonance imaging and speed of mental processing. *Archives of neurology*, *47*(2), 151–156.

Kapeller, P., Barber, R., Vermeulen, R. J., Adèr, H. J., Scheltens, P., Freidl, W., Almkvist, O., et al. (2003). Visual rating of age-related white matter changes on magnetic resonance imaging: scale comparison, interrater agreement, and correlations with quantitative measurements. *Stroke; a journal of cerebral circulation*, *34*(2), 441–445.

Lamar, M., Price, C. C., Libon, D. J., Penney, D. L., Kaplan, E., Grossman, M., & Heilman, K. M. (2007). Alterations in working memory as a function of leukoaraiosis in dementia. *Neuropsychologia*, *45*(2), 245–254. doi:10.1016/j.neuropsychologia.2006.07.009

Lamar, M., Catani, M., Price, C. C., Heilman, K. M., & Libon, D. J. (2008). The impact of region-specific leukoaraiosis on working memory deficits in dementia. *Neuropsychologia*, *46*(10), 2597–2601. doi:10.1016/j.neuropsychologia.2008.04.007

Lao, Z., Shen, D., Liu, D., Jawad, A., Melhem, E. R., Launer, L., Bryan, R., et al. (2008). Computer-Assisted Segmentation of White Matter Lesions in 3D MR Images Using Support Vector Machine1. *Academic radiology*, *15*(3), 300–313. doi:10.1016/j.acra.2007.10.012

Liao, D., Cooper, L., Cai, J., Toole, J. F., Bryan, N. R., Hutchinson, R. G., & Tyroler, H. A. (1996). Presence and severity of cerebral white matter lesions and hypertension, its treatment, and its control. The ARIC Study. Atherosclerosis Risk in Communities Study. *Stroke; a journal of cerebral circulation*, *27*(12), 2262–2270.

Libon, D. J., Bogdanoff, B., Cloud, B. S., Skalina, S., Giovannetti, T., Gitlin, H. L., & Bonavita, J. (1998). Declarative and Procedural Learning, Quantitative Measures of the Hippocampus, and Subcortical White Alterations in Alzheimer's Disease and Ischaemic Vascular Dementia. *Journal of Clinical and Experimental Neuropsychology*, *20*(1), 30–41. doi:info:doi/10.1076/jcen.20.1.30.1490

Libon, D. J., Price, C. C., Giovannetti, T., Swenson, R., Bettcher, B. M., Heilman, K. M., & Pennisi, A. (2008). Linking MRI hyperintensities with patterns of neuropsychological impairment: evidence for a threshold effect. *Stroke; a journal of cerebral circulation*, *39*(3), 806–813. doi:10.1161/STROKEAHA.107.489997

Libon, D. J., Price, C. C., Davis Garrett, K., & Giovannetti, T. (2004). From Binswanger's disease to leuokoaraiosis: what we have learned about subcortical vascular dementia. *The Clinical Neuropsychologist*, *18*(1), 83–100.

Lindgren, A., Roijer, A., Rudling, O., Norrving, B., Larsson, E. M., Eskilsson, J., Wallin, L., et al. (1994). Cerebral lesions on magnetic resonance imaging, heart disease, and vascular risk factors in subjects without stroke. A population-based study. *Stroke; a journal of cerebral circulation*, *25*(5), 929–934.

Mahon, K., Wu, J., Malhotra, A. K., Burdick, K. E., DeRosse, P., Ardekani, B. A., & Szeszko, P. R. (2009). A voxel-based diffusion tensor imaging study of white matter in bipolar disorder. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, *34*(6), 1590–1600. doi:10.1038/npp.2008.216

Maillard, P., Crivello, F., Dufouil, C., Tzourio-Mazoyer, N., Tzourio, C., & Mazoyer, B. (2009). Longitudinal follow-up of individual white matter hyperintensities in a large cohort of elderly. *Neuroradiology*, *51*(4), 209–220. doi:10.1007/s00234-008-0489-0

Manolio, T. A., Kronmal, R. A., Burke, G. L., Poirier, V., O'Leary, D. H., Gardin, J. M., Fried, L. P., et al. (1994). Magnetic resonance abnormalities and cardiovascular disease in older adults. The Cardiovascular Health Study. *Stroke; a journal of cerebral circulation*, *25*(2), 318–327.

McGraw, K., & Wong, S. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, *1*(1), 30–46.

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., et al. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage*, *40*(2), 570–582. doi:10.1016/j.neuroimage.2007.12.035

Mori, S., & van Zijl, P. (2007). Human white matter atlas. *American Journal of Psychiatry*, *164*(7), 1005. doi:10.1176/appi.ajp.164.7.1005

Munoz, D. G., Hastak, S. M., Harper, B., Lee, D., & Hachinski, V. C. (1993). Pathologic correlates of increased signals of the centrum ovale on magnetic resonance imaging. *Archives of neurology*, *50*(5), 492–497.

Nguyen, T. H., Yoshida, M., Stievenart, J. L., Iba-Zizen, M. T., Bellinger, L., Abanou, A., Kitahara, K., et al. (2005). MR tractography with diffusion tensor imaging in clinical routine. *Neuroradiology*, *47*(5), 334–343. doi:10.1007/s00234-005-1338-z

Pantoni, L., & Garcia, J. H. (1995). The significance of cerebral white matter abnormalities 100 years after Binswanger's report. A review. *Stroke; a journal of cerebral circulation*, *26*(7), 1293–1301.

Pantoni, L., & Garcia, J. H. (1997). Pathogenesis of leukoaraiosis: a review. *Stroke; a journal of cerebral circulation*, *28*(3), 652–659.

Price, C. C., Jefferson, A. L., Merino, J. G., Heilman, K. M., & Libon, D. J. (2005a). Subcortical vascular dementia: integrating neuropsychological and neuroradiologic data. *Neurology*, *65*(3), 376–382. doi:10.1212/01.wnl.0000168877.06011.15

Price, C. C., Schmalfuss, I., & Sistrom, C. (2005b). Quantification of white matter alterations: A reliability analysis. *Journal of the International Neuropsychological Society: JINS*, *11*(S1), 115–116.

Prins, N. D., van Straaten, E. C. W., van Dijk, E. J., Simoni, M., van Schijndel, R. A., Vrooman, H. A., Koudstaal, P. J., et al. (2004). Measuring progression of cerebral white matter lesions on MRI: visual rating and volumetrics. *Neurology*, *62*(9), 1533–1539.

Rasband, W. S. (1997). ImageJ. Bethesda, MD: U. S. National Institutes of Health. Retrieved from http://imagej.nih.gov/ij/

Román, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H., Amaducci, L., et al. (1993). Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology*, *43*(2), 250–260.

Rydberg, J. N., Hammond, C. A., Grimm, R. C., Erickson, B. J., Jack, C. R., Huston, J., & Riederer, S. J. (1994). Initial clinical experience in MR imaging of the brain with a fast fluid-attenuated inversion-recovery pulse sequence. *Radiology*, *193*(1), 173–180.

Sachdev, P. S., Parslow, R., Wen, W., Anstey, K. J., & Easteal, S. (2009). Sex differences in the causes and consequences of white matter hyperintensities. *Neurobiology of aging*, *30*(6), 946–956. doi:10.1016/j.neurobiolaging.2007.08.023

Sachdev, P. S., Wen, W., Chen, X., & Brodaty, H. (2007). Progression of white matter hyperintensities in elderly individuals over 3 years. *Neurology*, *68*(3), 214–222. doi:10.1212/01.wnl.0000251302.55202.73

Scarpelli, M., Salvolini, U., Diamanti, L., Montironi, R., Chiaromoni, L., & Maricotti, M. (1994). MRI and pathological examination of post-mortem brains: the problem of white matter high signal areas. *Neuroradiology*, *36*(5), 393–398.

Scheltens, P., Barkhof, F., Leys, D., Pruvo, J. P., Nauta, J. J., Vermersch, P., Steinling, M., et al. (1993). A semiquantative rating scale for the assessment of signal hyperintensities on magnetic resonance imaging. *J Neurol Sci*, *114*(1), 7–12.

Schmidt, R., Fazekas, F., Kleinert, G., Offenbacher, H., Gindl, K., Payer, F., Freidl, W., et al. (1992). Magnetic resonance imaging signal hyperintensities in the deep and subcortical white matter. A comparative study between stroke patients and normal volunteers. *Archives of neurology*, *49*(8), 825–827.

Schmidt, R., Fazekas, F., Offenbacher, H., Dusek, T., Zach, E., Reinhart, B., Grieshofer, P., et al. (1993). Neuropsychologic correlates of MRI white matter hyperintensities: a study of 150 normal volunteers. *Neurology*, *43*(12), 2490–2494.

Schmidt, R., Fazekas, F., Kapeller, P., Schmidt, H., & Hartung, H. P. (1999). MRI white matter hyperintensities: three-year follow-up of the Austrian Stroke Prevention Study. *Neurology*, *53*(1), 132–139.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, *86*(2), 420–428.

Silbert, L. C., Howieson, D. B., Dodge, H., & Kaye, J. A. (2009). Cognitive impairment risk: white matter hyperintensity progression matters. *Neurology*, *73*(2), 120–125. doi:10.1212/WNL.0b013e3181ad53fd

Silbert, L. C., Nelson, C., Howieson, D. B., Moore, M. M., & Kaye, J. A. (2008). Impact of white matter hyperintensity volume progression on rate of cognitive and motor decline. *Neurology*, *71*(2), 108–113. doi:10.1212/01.wnl.0000316799.86917.37

Skoog, I., Marcusson, J., & Blennow, K. (1998). Dementia. It's getting better all the time. *Lancet*, *352 Suppl 4*, SIV4.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl 1*, S208–19. doi:10.1016/j.neuroimage.2004.07.051

Sørensen, T. (1957). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biologiske Skrifter / Kongelige Danske Videnskabernes Selskab*, *5*, 1–34.

Tate, D. F., Jefferson, A. L., Brickman, A. M., Hoth, K. F., Gunstad, J., Bramley, K., Paul, R. H., et al. (2008). Regional White Matter Signal Abnormalities and Cognitive Correlates Among Geriatric Patients with Treated Cardiovascular Disease. *Brain Imaging and Behavior*, *2*(3), 200–206. doi:10.1007/s11682-008-9032-5

Tomimoto, H., Akiguchi, I., Suenaga, T., Nishimura, M., Wakita, H., Nakamura, S., & Kimura, J. (1996). Alterations of the blood-brain barrier and glial cells in white-matter lesions in cerebrovascular and Alzheimer's disease patients. *Stroke; a journal of cerebral circulation*, *27*(11), 2069–2074.

van Straaten, E. C. W., Fazekas, F., Rostrup, E., Scheltens, P., Schmidt, R., Pantoni, L., Inzitari, D., et al. (2006). Impact of white matter hyperintensities scoring method on correlations with clinical data: the LADIS study. *Stroke; a journal of cerebral circulation*, *37*(3), 836–840. doi:10.1161/01.STR.0000202585.26325.74

Vermeer, S. E., van Dijk, E. J., Koudstaal, P. J., Oudkerk, M., Hofman, A., Clarke, R., & Breteler, M. M. B. (2002). Homocysteine, silent brain infarcts, and white matter lesions: The Rotterdam Scan Study. *Annals of Neurology*, *51*(3), 285–289.

Wahlund, L., Barkhof, F., Fazekas, F., Bronge, L., Augustin, M., Sjögren, M., Wallin, A., et al. (2001). A new rating scale for age-related white matter changes applicable to MRI and CT. *Stroke; a journal of cerebral circulation*, *32*(6), 1318–1322.

Wen, W., & Sachdev, P. S. (2004a). The topography of white matter hyperintensities on brain MRI in healthy 60- to 64-year-old individuals. *NeuroImage*, *22*(1), 144–154. doi:10.1016/j.neuroimage.2003.12.027

Wen, W., & Sachdev, P. S. (2004b). Extent and distribution of white matter hyperintensities in stroke patients: the Sydney Stroke Study. *Stroke; a journal of cerebral circulation*, *35*(12), 2813–2819. doi:10.1161/01.STR.0000147034.25760.3d

Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T. E. J., Beckmann, C. F., et al. (2009). Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, *45*(1 Suppl), S173–86. doi:10.1016/j.neuroimage.2008.10.055

Wu, M., Rosano, C., Butters, M., Whyte, E., Nable, M., Crooks, R., Meltzer, C. C., et al. (2006). A fully automated method for quantifying and localizing white matter hyperintensities on MR images. *Psychiatry research*, *148*(2-3), 133–142. doi:10.1016/j.pscychresns.2006.09.003

Ylikoski, R., Ylikoski, A., Erkinjuntti, T., Sulkava, R., Raininko, R., & Tilvis, R. (1993). White matter changes in healthy elderly persons correlate with attention and speed of mental processing. *Archives of neurology*, *50*(8), 818–824.

Young, V. G., Halliday, G. M., & Kril, J. J. (2008). Neuropathologic correlates of white matter hyperintensities. *Neurology*, *71*(11), 804–811. doi:10.1212/01.wnl.0000319691.50117.54

Zijdenbos, A. P., Dawant, B. M., Margolin, R. A., & Palmer, A. C. (1994). Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE transactions on medical imaging*, *13*(4), 716–724. doi:10.1109/42.363096

## BIOGRAPHICAL SKETCH

Stephen Towler is primarily interested in using human brain imaging to investigate the neural correlates of typical development, aging, neurodegenerative disease, debilitating psychiatric illness, and individual differences in language and cognition. He has designed and executed all stages of human brain structural, functional, and diffusion magnetic resonance imaging studies for National Institutes of Health sponsored research, and has enjoyed working directly with healthy volunteers, as well as a range of adult and child patients. His current methodological focus is the sensitivity and reliability of functional magnetic resonance imaging in pre/post-treatment assessment of stroke patients. He graduated from the University of Florida with a Bachelor of Science in Psychology in 2008, and from the Florida School for Massage in 2010. In his future therapeutic work, he is looking forward to exploring the integration of