A Comparison of Qualitative and Quantitative White Matter Methods in Pediatric Traumatic Brain Injury

Kacie LaRae Wright

Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

BYU ScholarsArchive Citation
Wright, Kacie LaRae, "A Comparison of Qualitative and Quantitative White Matter Methods in Pediatric Traumatic Brain Injury" (2018). All Theses and Dissertations. 7384.
https://scholarsarchive.byu.edu/etd/7384

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
A Comparison of Qualitative and Quantitative White Matter Methods in Pediatric Traumatic Brain Injury

Kacie LaRae Wright

A dissertation submitted to the faculty of Brigham Young University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Ramona O. Hopkins, Chair
   Erin D. Bigler
   C. Brock Kirwan
   Shawn D. Gale
   Michael J. Larson

Department of Psychology
Brigham Young University

Copyright © 2018 Kacie LaRae Wright
All Rights Reserved
ABSTRACT

A Comparison of Qualitative and Quantitative White Matter Methods in Pediatric Traumatic Brain Injury

Kacie LaRae Wright
Department of Psychology, BYU
Doctor of Philosophy

Magnetic resonance imaging is used to assess white matter (WM) abnormalities including total WM volumes and WM hyperintensities (WMHs). Comparisons between several qualitative and quantitative methods to assess WM that are used in research and clinical settings are lacking in pediatric traumatic brain injury (TBI). This study 1) WM methods including Scheltens ratings, manual tracings, NeuroQuant®, and FreeSurfer; (2) compared total WM volumes and WMHs to assess potentially similar reporting of WM integrity; and (3) assessed the relationship between cognitive functions (intelligence, attention, processing speed, and language) and WM in pediatric TBI.

Sixty participants (65% male) between ages 8-13 years old, had a complicated-mild (53%), moderate (15%), or severe TBI (32%) with a mean age of 2.6 at the time of injury. NeuroQuant® WMH volumes had poor agreement ($ICC = .24$), and did not correlate ($r = .12, p = .21$) with manually traced WMH volumes. Scheltens WMH ratings had good to excellent agreement and correlated with NeuroQuant® ($ICC = .62; r = .29, p = .005$) and manually traced WMH volumes ($ICC = .82; r = .50, p = .000$). NeuroQuant® and FreeSurfer total WM volumes had fair agreement and were correlated ($ICC = .52; r = .38, p = .004$). No significant difference in total WM volumes were found between complicated-mild and moderate-severe TBI groups, and in subgroups with and without WMHs. Processing speed was significantly associated with Scheltens WMH ratings: $p = .004$, manually traced WMHs: $p = .002$, and NeuroQuant® WMHs: $p = .007$. No other association between cognitive functions and WM volumes or hyperintensities were found. Correlations between NeuroQuant® and manual tracings with processing speed differed by sex, where males had significant correlations but females did not.

Deciding when to use manual tracing and NeuroQuant® WMH volumes and Scheltens ratings in clinical or research settings will depend on available resources (e.g., time, technology, funding, and expertise) and purpose of assessing WMHs. Total WM volumes did not appear to capture WM pathology as assessed by WMHs, likely due to the sample being underpowered and that total WM volumes possibly included WMHs. Limitations include restricted range of injury severity, heterogeneity of lesions, and small sample size. Additional research is needed in a larger sample of pediatric TBI.

Keywords: white matter hyperintensities, pediatric, volumes, ratings, traumatic brain injury
ACKNOWLEDGEMENTS

This dissertation would not have been possible without the Social Outcomes in Brain Injury Kids (SOBIK) dataset, provided by Dr. Erin D. Bigler and the SOBIK investigators. I am deeply indebted to the countless individuals who helped fund, recruit, collect MRI images, administer neuropsychological testing, and otherwise contribute to this impressive dataset. I am also grateful to Dr. Erin D. Bigler, Dr. Ramona O. Hopkins, and the Department of Psychology for their financial support, which allowed me to collect manual and NeuroQuant® reports for all of my participants.

I am grateful to all those with whom I have had the pleasure of working in completing the dissertation, as well as other projects. Each member of my committee has contributed much to the overall development of this dissertation. I would especially like to thank my mentor Dr. Ramona Hopkins, the chair of my committee, without whose faith, strength, and guidance I would not be who I am today. She has taught me through her example what a good scientist, instructor, and person should be. I am also grateful to Tracy Abildskov, Dr. Naomi Goodrich-Hunsaker, Dr. C. Brock Kirwan, and Dr. Derin Cobia for sharing their knowledge and expertise in neuroimaging analyses, and Dr. Joseph Olsen for his statistical insights, which provided guidance and clarity. In addition, I would like to thank Frank Roberson, an undergraduate research assistant, for his tireless efforts in helping complete the manual tracings with me.

Finally, the support and encouragement I have received from my wonderful family and friends was essential to my success. Their unwavering love and support have been pillars of strength. Most importantly, I wish to thank my loving husband, Gart, who has provided unending support, encouragement, and inspiration.
TABLE OF CONTENTS

ABSTRACT....................................................................................................................................... ii

ACKNOWLEDGEMENTS............................................................................................................. iii

LIST OF TABLES....................................................................................................................... viii

LIST OF FIGURES .................................................................................................................... ix

A Comparison of Qualitative and Quantitative White Matter Methods in Pediatric Traumatic Brain Injury ................................................................................................................................. 1

Pediatric Traumatic Brain Injury ............................................................................................... 2

Developmental Factors and Pediatric Traumatic Brain Injury .................................................... 4

Mechanisms of Injury ................................................................................................................. 5

Classification of Injury Severity ............................................................................................... 6

Neuroimaging ............................................................................................................................... 8

White Matter Atrophy ................................................................................................................. 10

White Matter Hyperintensities (WMHs) .................................................................................. 11

Incorporating Quantitative Neuroimaging into Clinical Practice .............................................. 14

Magnetic Resonance Imaging Methods for White Matter Assessment ........................................ 16

Scheltens Rating Scale .............................................................................................................. 18

Manually Traced Brain Volumes by an Expert Rater ................................................................... 19

Automated Segmentation Methods .......................................................................................... 20

Comparison of Total WM and WMH Methods ......................................................................... 24

Cognitive Function ....................................................................................................................... 24
Processing Speed.................................................................................................................................................26

Intellectual Function..............................................................................................................................................27

Attention..........................................................................................................................................................29

Language..........................................................................................................................................................30

Sex Differences ...............................................................................................................................................31

Purpose.............................................................................................................................................................31

Aims and Hypotheses ...........................................................................................................................................32

Aim 1: Comparison of Quantitative White Matter Hyperintensity Methods ..................................................32

Aim 2: Comparison of Qualitative and Quantitative White Matter Hyperintensity Methods .....................32

Aim 3: Comparison of Quantitative Total White Matter Methods ...................................................................33

Aim 4: Comparison of Total White Matter Volumes and White Matter Hyperintensities .........................33

Research Question: Comparison of Cognitive Function and Methods of Assessing White Matter ..................34

Methods ..........................................................................................................................................................35

Participants ......................................................................................................................................................35

Neuropsychological Tests ................................................................................................................................37

Magnetic Resonance Imaging ............................................................................................................................39

Qualitative Rating of White Matter Hyperintensities ....................................................................................42

Volume Acquisition ..........................................................................................................................................43

Statistical Analysis ............................................................................................................................................45

Missing Data .....................................................................................................................................................45

Statistical Analyses for Normality ....................................................................................................................46

Comparison of Quantitative White Matter Hyperintensity Methods ............................................................46
Comparison of Total White Matter Volumes and White Matter Hyperintensities .......... 74

Research Question: Comparison of Cognitive Function and Methods of Assessing White Matter .................................................................................................................................................. 75

Sex Differences ........................................................................................................................................ 78

Study Strengths and Limitations .................................................................................................................. 79

Conclusions ..................................................................................................................................... 81

References..................................................................................................................................... 83

Appendix A: Example of NeuroQuant®’s Multi Structure Atrophy Report ......................... 126

Appendix B: Example of NeuroQuant®’s LesionQuant Report .............................................. 129

Appendix C: Description of Neuropsychological Instruments ..................................................... 132

Appendix D: Full Scheltens Rating Scale ..................................................................................... 135

Appendix E: Example of Scheltens Ratings for White Matter Regions ................................... 136

Appendix F: Log Transformed Correlation between Manual Tracing and NeuroQuant® .......... 137

Appendix G: Summary of Neuroimaging Data with Outliers Excluded ................................... 138
LIST OF TABLES

Table 1. White Matter Methods Used in the Current Study ..........................................................18
Table 2. Complete List of Cognitive Tests Obtained in the SOBIK dataset and Used in the Current
Study ...................................................................................................................................... 38
Table 3. Modified Scheltens Rating Scale for White Matter Hyperintensities ......................... 43
Table 4. Demographic and Medical Data of TBI Children .......................................................... 53
Table 5. Ratings or Volumetric Data Adjusted for TICV Within Method, Grouped by Injury
Severity and Total Group ...................................................................................................... 54
Table 6. Neuropsychological Data Grouped by Injury Severity and Total Group ....................... 55
Table 7. Review of WMH Method Comparison Findings .......................................................... 59
Table 8. Strength and Significance of Correlations between White Matter Methods and Cognitive
Measures in SEM Model ...................................................................................................... 62
Table 9. Chi-square Difference Test for Sex Differences for the Model and Individual
Comparisons between WM Methods and Cognitive Functions, Regular Analysis and
Outliers Excluded .................................................................................................................. 63
Table 10. Volumetric Data for Each Method with Outliers Excluded, Grouped by Injury Severity
and Total Group ................................................................................................................... 139
Table 11. Correlation and Inter Class Correlations (ICC) with Transformations with Outliers
Removed ...................................................................................................................................... 141
LIST OF FIGURES

Figure 1. Example of FLAIR white matter hyperintensities (WMHs) on an axial scan in a child who experienced a severe TBI and sustained a frontal contusion ........................................... 13

Figure 2. Example of hypointensities and hyperintensities on T1, FLAIR, and T2 images in an individual with pediatric TBI ................................................................................................ 14

Figure 3. Example of ROI segmentation using a static and dynamic atlas ............................................. 23

Figure 4. Diagram summarizing image sequences used for each WM neuroimaging methods .................................................. 41

Figure 5. A visual of hypotheses in Aims 1 and 2 ................................................................................. 47

Figure 6. A visual of hypothesis in Aims 3 .......................................................................................... 49

Figure 7. Visual of hypothesis 4a. Comparing total WM volumes between complicated-mild and moderate-severe TBI participants .............................................................. 49

Figure 8. Visual of hypotheses 4b: Comparing total WM volumes in those with and without WMHs in both complicated-mild and moderate-severe TBI participants ........................................ 50

Figure 9. Visual of the research question: All variables are co-varied. The multiple group analysis assessed sex differences and the associations between WM methods and cognitive functions ........................................... 51

Figure 10. For WM method comparisons, correlational scatterplots are to the left and Bland-Altman plots assessing agreement are to the right ......................................................... 57

Figure 11. Boxplot of total WM volumes for participants with and without WMHs in complicated-mild and moderate-severe TBI groups ................................................................. 61

Figure 12. Descriptive summary of methodological differences for identifying WMHs between methods .................................................................................................................. 69
Figure 13. A search of PubMed using search terms: “NeuroQuant”, “LesionQuant”, “Scheltens rating”, “manual tracing”, and “FreeSurfer” was conducted................................................ 71

Figure 14. Example of the Multi Structural Atrophy Report from NeuroQuant®, which used T1 images to quantify whole brain, lateral ventricles, thalamus, cortical gray matter, cerebral white matter, white matter hypointensities, third ventricle, hippocampi, and inferior lateral ventricle volumes................................................................................................................. 128

Figure 15. Example of NeuroQuant® ’s LesionQuant report, which used FLAIR images to quantify white matter hyperintensities (WMHs) in various brain regions. Anatomical distribution of the hyperintensities are assessed, as well as the number of lesions, the volume, the corrected volume for total intracranial volume, and lesion burden are calculated ............................................................................................................................................. 131

Figure 16. Visual of the full Scheltens rating scale used to assess hyperintensities on T2 images; The modified scale for assessing WMHs excludes all GM regions ......................... 135

Figure 17. Example of Scheltens ratings for white matter hyperintensities in pediatric TBI participants........................................................................................................................... 136

Figure 18. Log transformed correlation between manual tracing and NeuroQuant® WMH volumes........................................................................................................................................................................ 137

Figure 19. Scatterplot of WM methods by injury severity (GCS)............................................ 139

Figure 20. Scatterplots with regression line for each WMH comparison with outliers excluded ............................................................................................................................................................................................................. 140
A Comparison of Qualitative and Quantitative White Matter Methods in Pediatric Traumatic Brain Injury

Pediatric traumatic brain injury (TBI) is a leading source of death and acquired disability in children and can abruptly change the lives of both the child and their family (Au & Clark, 2017; Babikian & Asarnow, 2009; Pinto, Poretti, Meoded, Tekes, & Huisman, 2012; National Institute of Neurological Disorders and Stroke [NINDS], 2015). By definition, TBI is an event that causes brain pathology and altered brain function following an external force to the head (Menon, Schwab, Wright, & Maas, 2010). Because external forces induce movement of brain tissue within the cranium, multiple physical factors have the potential to injure the brain. For example, external force can cause deformation throughout the brain with the potential for surface areas to collide with the dura, and rough boney prominences of the inner skull (e.g., sphenoid ridge; Armstrong, Mierzwa, Marion, & Sullivan, 2016; Bigler, 2007; Wright & Ramesh, 2012). Mechanical rotational and translational forces, as well as shock waves caused by skull vibrations from the impact can damage gray matter (GM) and white matter (WM) due to differences in tissue elasticity (Frati et al., 2017; Martin, 2016; Wright & Ramesh, 2012). The stretching and twisting of brain tissue results in damage to the microvascular and neuronal structures including shearing injuries that result in diffuse axonal injury (DAI; Frati et al., 2017; Martin, 2016; de la Plata et al., 2007). Accordingly, TBI is often the source of considerable WM damage, often viewed as a general traumatic axonal injury or TAI (Buki & Povlishock, 2006; Hill, Coleman, & Menon, 2016) which can be detected by neuroimaging (Armstrong et al., 2016; Jarrett et al., 2016; Lee & Newberg, 2005). Depending on the location and severity of injury, TBI and its associated structural damage can result in significant impairments in cognitive (Catroppa, Anderson, Beauchamp, & Yeates, 2016; Lloyd, Wilson, Tenovuo, & Saarijarvi, 2015), social
COMPARISON OF WM METHODS IN PEDIATRIC TBI

(Rigon, Voss, Turkstra, Mutlu, & Duff, 2016; Yeates et al., 2013), emotional (van der Horn, Liemburg, Aleman, Spikman, & van der Naalt, 2016), and behavioral functioning (Königs et al., 2018).

Advancements in neuroimaging such as magnetic resonance imaging (MRI), allow for investigations of brain structures regions and networks (Bigler, 2007; Bigler, 2015; Bigler, 2017; Murphy & Duhaime, 2016) along with their associated cognitive and behavioral functions (Araujo et al., 2017; Yeates et al., 2013). Neuroimaging has become especially important in understanding TBI as it provides insight into the severity and extent of brain pathology, and can aid in diagnosis, prognosis, treatment, and evaluation of recovery and rehabilitation (Bigler, 2015; Bigler, 2017; Brewer, 2009; Jantz & Bigler, 2014; Murphy & Duhaime, 2016; Tieves, 2010). Since current magnetic resonance imaging (MRI) methods permit the examination of WM integrity, this study uses neuroimaging methods to assess WM abnormalities in children who have experienced a TBI (Martin, 2016; McAllister et al., 2012).

Pediatric Traumatic Brain Injury

TBI affects approximately 750 per 100,000 children each year in the United States (Anderson, Godfrey, Rosenfeld, & Catroppa, 2012). Individuals who experience TBI may recover without significant sequelae, especially individuals with mild TBI, or may develop severe and persistent long-term morbidities, depending on type, location, and severity of the injury (Anderson et al., 2012; Arroyos-Jurado, Paulsen, Ehly, & Max, 2006; Au & Clark, 2017). Because of its persistent and often permanent impairments in physical, cognitive, social, and psychological function, especially for individuals with moderate to severe TBI, TBI meets the criteria of the World Health Organization (WHO) definition of chronic diseases (Masel, 2015; Masel & DeWitt, 2010; World Health Organization [WHO], 2002). TBI is often referred to as a
silent injury, as many individuals lack obvious external signs of their injuries such as paralysis or speech impairments, but still experience less obvious, but no less significant cognitive impairments and behavioral changes (Bigler, 2003; Catroppa et al., 2016; Faul, Xu, Wald, & Coronado, 2010; Langlois, Marr, Mitchko, & Johnson, 2005). This is especially true following mild TBI, but also occurs after moderate and severe TBI (Faul et al., 2010; Langlois et al., 2005). Because of the absence of external manifestations of the TBI, the number of children and adults with permanent disability due to TBI likely is under recognized and under reported, as many individuals do not seek medical treatment or professional assistance following their injury (Langlois, Rutland-Brown, & Thomas, 2004; NINDS, 2015).

The etiologies of pediatric TBI are variable and include motor vehicle accidents, sports and recreational injuries, falls, assaults, or being hit in the head by an object (Centers for Disease Control and Prevention [CDC], 2016; Faul et al., 2010; NINDS, 2015). In children under 14 years of age, falls are the most common cause of TBI (55%), with accidental blunt trauma (e.g., sport related injuries) and motor vehicle accidents cause almost a quarter of all TBIs and TBI-related deaths (CDC, 2016; Faul et al., 2010; NINDS, 2015). Unfortunately, pediatric TBI can also occur as a result of parental/caregiver violence (e.g., shaken baby syndrome) or neglect (e.g., failure to use proper safety devices in vehicles; Hooper et al., 2004; NINDS, 2015). Males are two to three times more likely than females to experience a TBI, as males are more likely to participate in high risk activities (e.g., riding a bike or scooter without a helmet, driving without wearing a seat belt) at rates higher than females (Eaton et al., 2010; Jantz, Davies, & Bigler, 2014).
Developmental Factors and Pediatric Traumatic Brain Injury

TBI can occur at any age however, developmental factors may put children at a unique risk for brain pathology, disrupting brain development (Arca, 2013; Bigler, 2007; Case, 2008; Pinto et al., 2012). Childhood is a period of rapid brain and cognitive growth (Toga, Thompson, & Sowell, 2006) with intracranial volume reaching 80% of the adult volumes by age two (Arca, 2013; Tieves, 2010), and 90-95% by age five (Caviness, Kennedy, Richelme, Rademacher, & Filipek, 1996; Courchesne et al., 2000). The expanding brain volume stimulates skull development, with the cranial vault reaching approximate adult size around age seven (Wolf et al., 2003). As a result of this interaction between brain development and intracranial expansion, a child’s brain fits very tightly in the cranial vault, resulting in a smaller subarachnoid space compared to that observed in the mature brain (Arca, 2013; Bigler, 2007; Case, 2008). The smaller subarachnoid space provides less buoyancy due to the smaller volume of cerebral spinal fluid, resulting in both potential protective as well as vulnerability factors. For example, a tight fit of the brain in the cranial vault may mitigate against movement, reducing injury effects on the mild end of the spectrum; alternatively, reduced protection or cushioning of the brain parenchyma due to less cerebral spinal fluid may contribute to diffuse cerebral injuries during trauma, especially contusions (Arca, 2013; Tieves, 2010).

Studies comparing traumatic axonal damage in myelinated and unmyelinated axons show that myelination may protect axons from trauma by making the axons more tolerant to some aspects of mechanical deformations (Johnson, Stewart, & Smith, 2013b; Reeves, Phillips, Lee, & Povlishock, 2007; Reeves, Phillips, & Povlishock, 2005; Staal & Vickers, 2011). Children are not born with extensive myelination, although it rapidly emerges (Toga et al., 2006). It is not until mid to late adolescence that the brain approaches adult myelination and WM volumes
(Arca, 2013; Bigler, 2007; Nagy, Westerberg, & Klingberg, 2004; Toga et al., 2006). As such, immature myelination results in axons that are more vulnerable to shear injuries, and the ability of the brain to absorb trauma-related forces is reduced (Case, 2008; Laskowitz & Grant, 2016; Pinto et al., 2012). Head size and neck strength are additional contributors to brain vulnerability in TBI. Children less than 8 years have a large head to body ratio resulting in a head that is much heavier than the adult head in proportion to body mass (Case, 2008; Pinto et al., 2012). The larger head and weaker neck muscles put children at greater risk for injury as the neck is less able to withstand the acceleration forces during TBI (Case, 2008; Pinto et al., 2012). In summary, the smaller subarachnoid space, immature myelination and large head to body ratio may increase vulnerability for neural injury in children and may continue to be factors even in adolescents (Arca, 2013; Bigler, 2007; Case, 2008; Pinto et al., 2012).

**Mechanisms of Injury**

A number of biomechanical and biochemical mechanisms damage the brain following TBI (Hopkins, Tate, & Bigler, 2005; Park et al., 2014). As introduced above, WM is particularly susceptible to damage from DAI due to the axon’s limited tolerance for mechanical deformation (Armstrong et al., 2016; Besenski, 2002; Johnson et al., 2013b; de la Plata et al., 2007; NINDS, 2013; Wright & Ramesh, 2012). Axonal damage occurs due to stretch injury as well as rotation and axon shearing during the rapid brain acceleration and deceleration during impact (Johnson et al., 2013a; Martin, 2016; Wright & Ramesh, 2012). Smaller, unmyelinated axons found in the corpus callosum are more vulnerable to stretch and disconnection than myelinated axons, as stated previously myelin may be neuroprotective (Johnson et al., 2013b; Reeves et al., 2007). Similarly, shear forces also damage the microvasculature resulting in vascular injuries, including
micro-hemorrhages which can be used as neuro-radiological confirmation of DAI (Howarth, Blackwell, & Ono, 2016; Narayana, 2017; Riedy et al., 2016).

Secondary pathophysiological responses can continue to cause or exacerbate brain damage immediately or in the hours and days following TBI, worsening the extent and severity of the initial injury (Armstrong et al., 2016; Kochanek et al., 2000; Martin, 2016; Maxwell, 2015; Park et al., 2014). Damage caused by secondary mechanisms can include excessive release of excitatory neurotransmitters (such as glutamate), rapid calcium influx resulting in excitotoxicity, breakdown of the blood brain barrier, triggering apoptosis (programmed cell death) and necrosis, edema, inflammation, and changes in cerebrovascular function to name a few (Frati et al., 2017; Ichkova et al., 2017; Kochanek et al., 2000; Narayana, 2017; Villapol, Byrnes, & Symes, 2014; Witcher, Eiferman, & Godbout, 2015). Glutamate is an abundant excitatory amino acid; however, toxic levels following brain injury can directly damage neurons (Tieves, 2010; Villapol et al., 2014). An excessive influx of cellular calcium causes energy depletion and oxidative stress (Frati et al., 2017; Kochanek et al., 2000). A breakdown of the blood brain barrier can allow blood, toxins, and larger proteins to pass into the brain, contributing to inflammation and in severe cases, hypoxia (Ichkova et al., 2017; Tieves, 2010). The above factors can exacerbate neuronal injury inducing necrosis and apoptosis in brain tissue (Park et al., 2014). Regardless of the mechanisms, secondary axonal injury can cause or exacerbate TBI associated morbidities (Park, Bell, & Baker, 2008).

**Classification of Injury Severity**

Each TBI is unique. Classification of TBI severity provides a metric for prognostication and treatment and is essential in making comparisons (similarities and differences) across individuals and studies (Menon et al., 2010). A number of scaled indicators based on the
individuals’ level of consciousness and neurological responsiveness during the acute phase of injury, are used to classify injury severity (Howarth et al., 2016; Jones, 1979). The most widely used measure of injury severity is the Glasgow Coma Scale (GCS), which measures eye opening responses (e.g., in response to pain or voice), verbal responsiveness (e.g., confusion, orientation, incoherent words), and motor responses (e.g., movements to pain, flexion, obedient to commands) following injury (Teasdale & Jennett, 1974). The GCS’s summed score classifies TBI as mild (GCS score from 13-15), moderate (GCS score from 9-12), or severe (GCS score from 3-8; Teasdale & Jennett, 1974). Lower GCS scores, which indicate increased injury severity, is associated with longer coma duration, increased incidence and severity of cognitive impairments and psychological disorders (MRC CRASH Trial Collaborators, 2008; Teasdale et al., 2014), whereas individuals with the mildest injury severity (higher GCS scores) have the best prognosis (Filippi & Sanelli, 2016; Hawryluk & Manley, 2015; Königs et al., 2018).

Various studies have further classified mild TBI as complicated-mild TBI, where GCS is 13-15 and there must be confirmed pathology such as skull fracture and/or acute intracranial hemorrhage or identifiable edema on acute CT scans (Bigler et al., 2013a; Gorman et al., 2016; Roberts, Mathias, & Rose, 2016; Williams, Levin, & Eisenberg, 1990; Yeates et al., 2013). The use of complicated-mild TBI can increase specificity of TBI diagnosis, as not all individuals classified with mild TBI (the most common classification of TBI) have observable pathology and tend to recover quickly without complications, whereas complicated-mild TBI has visible brain pathology that is associated with an increased risk for slow or incomplete recovery (Iverson, 2006; Panenka et al., 2015; Veeramuthu et al., 2017).
Neuroimaging

Contemporary neuroimaging technology such as MRI allows clinicians and researchers to visually identify in vivo structural anatomy, including pathology (≥ 1mm³) that occurs following TBI (Bigler, 2015; Fisch, Konen, Halevy, Cohen, & Shuper, 2012; Le & Gean, 2009; Pinto et al., 2012). Neuroimaging can detect a variety of pathologies associated with TBI including skull fractures (Alcalá-Galiano et al., 2008; Roguski et al., 2015), focal lesions (Jarrett et al., 2016; Pinto et al., 2012), brain atrophy (Bigler et al., 2010; Zagorchev et al., 2016), ventricular enlargement (Dennis et al., 2016; Sherer & Sander, 2014), encephalomalacia (Bigler, 2015; Le & Gean, 2009), sulcal widening (Lee & Newberg, 2005; Pinto et al., 2012), along with GM and WM abnormalities (Bigler et al., 2013a; Jarrett et al., 2016; Lee & Newberg, 2005). Despite the heterogeneity of focal shear-related pathology observed on neuroimaging, some form of WM damage is the most widespread and consistent pathology following pediatric TBI (Armstrong et al., 2016; Johnson et al., 2013b; Tang-Schomer, Johnson, Baas, Stewart, & Smith, 2012).

Given contemporary image quantification techniques, WM pathology is frequently assessed using volumetric methods (Ferrer et al., 2013; Königs et al., 2018) and diffusion tensor imaging (DTI; Alexander, Lee, Lazar, & Field, 2007; Roberts et al., 2016; Wozniak et al., 2007). Volumes are considered a proxy for the health or integrity of a brain structure. The assumption is that when the structural volume of a region of interest (ROI) falls within some average range of healthy control volumes, or does not differ significantly from healthy control volumes, this indicates intact white matter (Albright, Leyden, & Airriess, 2016; Bigler et al., 2010; Brewer, 2009; Luo, Airriess, & Albright, 2015; Sidaros et al., 2009). Conversely, reduced total WM volume is a marker of atrophy indicating loss of brain integrity (Farbota et al., 2012; Königs et al., 2018). DTI is also an assessment of WM integrity, that measures the diffusion rates and
direction of water molecules along axons, where lower fractional anisotropy (FA) values signal areas of decreased WM integrity (axonal damage) via isotropic or non-directional diffusion of water (Adams et al., 2011; Alexander et al., 2007). The relationship between DTI and volumetric measures is complex and unclear as each method assesses very different components of WM. The focus of this paper is on comparing volumetric methods of total WM and WMHs, as such DTI and FA will not be assessed.

Differences in signal intensities within the WM may also indicate WM pathology. These regions of intense signal beyond what is observed in healthy WM have been referred to as “hyperintensities” or WM hyperintensities (WMHs), reflecting focal signs of WM pathology (Bigler et al., 2013a; Debette & Markus, 2010; Zhan et al., 2009). The presence of WMHs implicates axonal damage including damaged myelin, axonal shearing, or traumatic infarcts (Zhan et al., 2009) representing areas of irreversible damage including neuronal death, volume loss, and axonal degeneration (Warner et al., 2010). WMHs are present both in the acute and chronic phases of brain injury (Parkinson et al., 2002; Warner et al., 2010). While some axonal recovery may occur following the acute phase, (Sidaros et al., 2009; Warner et al., 2010), chronic axonal damage persists for months to years after brain injury, suggesting more permanent white matter damage (Kraus et al., 2007; Masel & DeWitt, 2010). Research suggests that some chronic WMHs may have filled with CSF, hence their hyperintense appearance in white matter on T2 images (Kim, MacFall, & Payne, 2008; Hernández et al., 2013). As such, WMHs may also reflect brain atrophy (Capizzano et al., 2004; Fujishima et al., 2014; Habes et al., 2016; Kochunov et al., 2010, Schmidt et al., 2005). The current study used MRI based methods to focus on WM integrity via assessing total WM volumes and WMHs in the chronic phase of pediatric TBI.
White Matter Atrophy

Following TBI non-viable neurons as well as macerated tissue are removed by macrophages, resulting in focal and diffuse brain atrophy (Ramlackhansingh et al., 2011; Witcher et al., 2015). If the damage is sufficiently extensive, the loss of neurons result in diffuse cerebral atrophy expressed as overall volume loss of brain parenchyma (Keightley et al., 2014b) most visibly evident as enlarged sulci (Bigler et al., 2010; Cole et al., 2018) and ventricles (Beauchamp et al., 2011; Dennis et al., 2016; Keightley et al., 2014b). Brain atrophy following TBI is due at least in part to WM damage (Cole et al., 2018; Gale, Johnson, Bigler, & Blatter, 1995; Johnson, Bigler, Burr, & Blatter, 1994). The corpus callosum, a midline structure that is especially sensitive to rotational forces, is one of the most vulnerable WM structures damaged by TBI (Bigler & Maxwell, 2011; Martin, 2016; McAllister et al., 2012). Damage to the corpus callosum includes focal WM lesions and atrophy (Martin, 2016). Atrophy of the corpus callosum occurs from generalized neuronal loss due to direct effects of the injury or by Wallerian degeneration of cortical neurons (Bigler, 2007; Johnson et al., 2013b). Since the corpus callosum is situated superior to and forms part of the ceiling boundary of the lateral ventricles along with the other non-collasal WM pathways that surround the ventricles, ventricle enlargement or hydrocephalus ex vacuo, is a passive reflection of reduced brain volume and loss of adjacent WM (Bigler et al., 2013a; Dennis et al., 2016). Reduction in total WM volume is common following TBI and is associated with ventricular enlargement (Beauchamp et al., 2011; Bigler et al., 2013a; Brezova et al., 2014; Königs et al., 2018). Since reduced total WM volumes, including the corpus callosum (Bigler et al., 2013a), occurs concurrently with increased ventricular volume (Bigler et al., 2013a; Zhou et al., 2013), total WM volume is a metric
commonly used to assess TBI outcome (Beauchamp et al., 2011; Bigler et al., 2013a; Königs et al., 2018; Zhou et al., 2013).

**White Matter Hyperintensities (WMHs)**

In healthy, typical developing children, WMHs are infrequent (Kim, Illes, Kaplan, Reiss, & Atlas, 2002). While WMHs occur in healthy older adults, especially after age 50 (Debette & Markus, 2010; Kim et al., 2008) and occasionally in young healthy adults (Hopkins et al., 2006), WMHs in children typically implicate some underlying pathologic change has occurred (Fisch et al., 2012). While WMHs are common following pediatric TBI (Bigler et al., 2013a), they are also observed in neuropsychiatric disorders such as bipolar disorder (Pillai et al., 2002), severe depression (Ehrlich et al., 2004) and schizophrenia (Alloza et al., 2016), as well as demyelinating disorders such as multiple sclerosis (Schmidt et al., 2012; Wang et al., 2016), vascular disease (Lambert et al., 2016; Righart et al., 2013), and normal aging (Habes et al., 2016; de Leeuw et al., 2001). Accordingly, the mere presence of a WMH is not diagnostic of a particular condition, but rather indicates some underlying pathological state.

WMHs indicate focal areas of WM pathology and manifest as bright, hyperintense signal on T2-weighted MRI sequences, especially the fluid attenuated inversion recovery (FLAIR) sequence (see Figure 1; Bigler et al., 2013a; Papp et al., 2014; Wu et al., 2006; Zhan et al., 2009). FLAIR images suppress cerebral spinal fluid (CSF) signal which separates the hyperintense signal abnormality from surrounding WM tissue with “normal” appearing signal, creating a distinct WMH boundary that can be quantified using manual or automated methods (Caligiuri et al., 2015; Luo et al., 2017b; de la Plata et al., 2007; Warner, de la Plata, Liebeskind, & Diaz-Arrastia, 2017). Although WMHs can be widely distributed throughout the brain following TBI (Bigler et al., 2013a), they frequently occur in the frontal and temporal regions.
(Bigler et al., 2013a; Cabana, Greenb, & Riedya, 2013). The fluid intense region adjacent to the ventricles (periventricular region) can also appear hyperintense on T2 images and may be erroneously included as pathology in WMH analyses (Alam & Sahu, 2010; Lisanti, Carlin, Banks, & Wang, 2007; Zimmerman, Fleming, Lee, Saint-Louis, & Deck, 1986). Dilated perivascular, or Virchow-Robin spaces, can be a sign of acute inflammation following TBI (Bigler & Maxwell, 2011; Inglese et al., 2005), and may present as hyperintense areas on standard T2 images but less so or absent on FLAIR imaging (see middle image on Figure 2). The differences in hyperintensities on T2 and FLAIR images help identify parameters used to define WMHs that may be clinically significant (Luo et al., 2017b). For the purposes of this investigation, WMH volumes were obtained using FLAIR images, where suspected hyperintensities were confirmed using T2 images for manual tracings (see methods for more details). WM hypointensities found on T1 images (Bigler, 2015; Sanfilipo, Benedict, Sharma, Weinstock-Guttman, & Bakshi, 2005) were not assessed in this study as they frequently underestimate or do not detect WM pathology identified by T2-based sequences (See Figure 2).
Figure 1. Example of FLAIR white matter hyperintensities (WMHs) on an axial scan in a child who experienced a severe TBI and sustained a frontal contusion. Hyperintense areas within the WM indicate WMHs (see arrow). As visualized anterior to the large WMH region, there is marked loss of brain parenchyma (dark area) with resulting encephalomalacia.
**Figure 2.** Example of hypointensities and hyperintensities on T1, FLAIR, and T2 images in an individual with pediatric TBI. The circle shows the location of the WM abnormality on all three images. The left image shows the T1 hypointensity or dark region which underestimates the extent of WM damage compared to the hyperintensities on the FLAIR (middle image) and T2 image (right image). Note the other punctate WM signal findings on the T2 image are perivascular or Virchow-Robin spaces.

**Incorporating Quantitative Neuroimaging into Clinical Practice**

Even though quantitative neuroimaging research has been ongoing for more than 30 years, there has been difficulty integrating quantitative neuroimaging findings in the clinical setting, especially because until recently all aspects of image quantification had to be done by manual tracing (Albright et al., 2016; Brewer, 2009). Quantitative neuroimaging provides objective methods to quantify pathology in research settings (Bigler, 2016; Ross, Graham, & Ochs, 2013) with the goal of providing empirical evidence to aid in clinical diagnosis, prognosis, management of symptoms, recovery, and help track individual progress over time, which may
lead to improved patient care (Albright & Roberts, 2015; Murphy & Duhaime, 2016). Though quantitative neuroimaging has clinical utility, it has not integrated well within the clinical workflow and currently remains predominately a research tool, especially in TBI (Brewer, 2009; Ross et al., 2013). There are several reasons for the lack of integration between research and clinical practice. Methods that use semi-automated or manual ROI segmentation require technology (e.g., Mac processors, access to sufficient computing power, expensive licensing) and high levels of training and expertise to obtain and interpret results (Brewer, 2009). Though more automated methods increase the ease of obtaining quantitative measures of the brain, such methods are not always practical for clinical use due to high costs, training, lack of universal availability, time to produce the results, and absence of agreed upon methods (Albright et al., 2016; Brewer, 2009).

Clinicians have long recognized the potential value of incorporating standardized reporting of neuroimaging findings, but to date clinical methods have largely been limited to visual estimates of WM pathology using qualitative and/or semi-quantitative rating scales (Bigler, 2016; Kim et al., 2008; Luo et al., 2017b; de la Plata et al., 2007). Some qualitative measures only report the presence or absence of WMHs, whereas other measures assess WMH occurrence, size, and location (Hopkins et al., 2006). Rating scales are useful and descriptive; however, do not capture subtle differences in WMH size nor provide quantitative volumes which may provide more exact and comparable information regarding the extent of the brain injury (Bigler, 2016; van den Heuvel et al., 2006b). While rating scales are used clinically and in research, rating scales may not be as reliable in tracking WMH changes over time (Caligiuri et al., 2015; Gouw et al., 2008; van den Heuvel et al., 2006b; Kapeller et al., 2003). To bridge the gap between clinical practicality and quantitative methods developed for research, automated
methods that are cost effective, quick, provide reliable volumetric data, and easy to interpret are essential to advance use in clinical settings (Albright et al., 2016; Brewer, 2009; Caligiuri et al., 2015).

**Magnetic Resonance Imaging Methods for White Matter Assessment**

There are a number of rating scales, manual, semi-automated, and automated programs used to assess WM integrity on neuroimaging. These programs or instruments include NeuroQuant® (www.cortechslabs.com), FreeSurfer (https://surfer.nmr.mgh.harvard.edu), Scheltens rating scale (Scheltens et al., 1993), manual tracings (Capizzano et al., 2004; Vannorsdall, Waldstein, Kraut, Pearlson, & Schretlen, 2009), Lesion Segmentation Tool (LST; http://www.applied-statistics.de/lst.html; Birdsill et al., 2014; Fujishima et al., 2014), or NeuroQuant®’s new LesionQuant report (Luo et al., 2017a; Luo et al., 2017b). Each program or instrument has merit and differs in terms of reliability, ease of use, and accessibility. Several programs such as FreeSurfer are open source and free to the public, whereas other programs such as LST in MATLAB (MATrix LABoratory; https://www.mathworks.com/products/matlab.html) require expensive licensing. Semi-automated programs are helpful but often require hands-on preprocessing of the MR images prior to automatic segmentation which are intensely operator-dependent, or requires the attention, time, and expertise of a trained user, which can affect the reliability and replicability of results (Brewer, 2009; Caligiuri et al., 2015).

Until recently, these programs have only been used in adult populations. Pediatrics is still a largely understudied population for neuroimaging analyses. This may be in part because of the difficulties of accounting for the rapid stages of development of pediatric brains. Because of the lasting effects of pediatric TBI on academic, social, and behavioral skills, it is important to assess WM integrity using neuroimaging programs in pediatric TBI population. The ability to use
neuroimaging methods to assess WM in pediatric TBI would allow clinicians and researchers the ability to link pathology to the cognitive and psychological outcomes.

Introduced about a decade ago, NeuroQuant®, is a fully automated, operator-independent program that allows for rapid ROI detection and volumetric assessment while minimizing user bias (Albright et al., 2016; Brewer, 2009; Fischl, 2012). NeuroQuant® is commercially available and marketed for clinical use (Tanpitukpongse, Mazurowski, Ikhena, & Petrella, 2017). NeuroQuant® has been used in adult populations including Alzheimer’s disease (Ochs, Ross, Zannoni, Abildskov, & Bigler, 2015; Persson et al., 2017; Tanpitukpongse et al., 2017), older adults with mild cognitive impairments (Wang et al., 2012), TBI (Ochs et al., 2015; Reid et al., 2017; Ross, Ochs, DeSmit, Seabaugh, & Havranek, 2015), dementia (Engedal, Braekhus, Andreassen, & Nakstad, 2012), multiple sclerosis (Beadnall et al., 2016; Luo et al., 2017b), chronic inflammation (McMahon, Shoemaker, & Ryan, 2016), and mesial temporal sclerosis (Azab, Carone, Ying, & Yousem, 2015), but its use has yet to be examined in pediatric populations including pediatric TBI.

Before NeuroQuant® can be used to assess pediatric TBI with confidence in research and/or clinical settings, its reliability with existing research and clinical methods must be determined (Koo & Li, 2016). In this case, NeuroQuant® results must be compared to a reliable and valid clinical rating scale (Scheltens), manual tracing the research gold standard for volumetric analysis and FreeSurfer a reliable automated, volumetric method that is widely used in research. Details of each of these methods are provided below. To date, only a handful studies have compared these imaging programs/instruments in adult populations, including TBI (Ross et al., 2015; Ochs et al., 2015), with most studies suggest good comparability between methods. NeuroQuant® brain volumes correlate well with neuroradiologist identification of global atrophy
COMPARISON OF WM METHODS IN PEDIATRIC TBI

(Ross et al., 2015), and FreeSurfer brain volumes (Ochs et al., 2015). Scheltens is correlated with manual volumetric analyses of WMHs in geriatrics (Gouw et al., 2008; van den Heuvel et al., 2006b; Kapeller et al., 2003). It is unclear if these tools will perform similarly in pediatric TBI in assessing WM integrity. To date, no study has investigated the interrelationships of these four different methods in pediatric TBI. Table 1 provides a list of MRI based WM methods and imaging sequences that are used in the current study.

Table 1

White Matter Methods Used in the Current Study

<table>
<thead>
<tr>
<th>Method</th>
<th>MRI Sequence</th>
<th>ROI Measurement</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheltens Rating</td>
<td>FLAIR</td>
<td>Hyperintensity ratings</td>
<td>total score (0 to 84)</td>
</tr>
<tr>
<td>(Scheltens et al., 1993)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manual Tracing *</td>
<td>FLAIR/T1/T2</td>
<td>WMH volumes</td>
<td>cm^3</td>
</tr>
<tr>
<td>NeuroQuant®</td>
<td>FLAIR/T1</td>
<td>WMH volumes</td>
<td>cm^3</td>
</tr>
<tr>
<td>(<a href="http://www.cortechslabs.com">www.cortechslabs.com</a>)</td>
<td></td>
<td>Total WM volumes</td>
<td>cm^3</td>
</tr>
<tr>
<td>FreeSurfer</td>
<td>T1</td>
<td>Total WM volumes</td>
<td>cm^3</td>
</tr>
<tr>
<td>(<a href="https://surfer.nmr.mgh.harvard.edu">https://surfer.nmr.mgh.harvard.edu</a>)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. ROI = region of interest; WMHs = white matter hyperintensities; FLAIR = fluid attenuated inversion recovery; * = see methods for more details.

Scheltens Rating Scale

The “gold standard” for assessing WMHs in a clinical setting are rating scales that either count or provides some index regarding the size and number of WMHs, with larger WMHs indicating increased severity (Scheltens et al., 1993; van Straaten et al., 2006). One of the most commonly used rating scales is the Scheltens rating scale (Scheltens et al., 1993). The Scheltens
rating scale is a semi-quantitative method that rates the presence, size, and location of WMHs on a T2, PD, or FLAIR image (Scheltens et al., 1993; van Straaten et al., 2006). The scores for each location are summed to create a total score that ranges from 0 to 84, with 84 signifying extensive hyperintensities throughout the brain (Scheltens et al., 1993; van Straaten et al., 2006). Scheltens rating scale assesses both gray matter and WMHs, but can easily be adapted to assess only WMHs, by summing only the WM regions. Rating scales and WMH volumes have been compared in the elderly (Gouw et al., 2008; van den Heuvel et al., 2006b; Kapeller et al., 2003) with good comparability; however, to date no studies have evaluated how such ratings relate to quantitative WMH volumes in pediatric TBI population.

**Manually Traced Brain Volumes by an Expert Rater**

Manually traced brain structures or regions can be used to provide detailed volumetric assessments of regions of interest (ROI; e.g., WMHs), and assess the accuracy of automated segmentation methods (Ambarki, Wahlin, Birgander, Eklund, & Malm, 2011; Grimm et al., 2015; Lyden et al., 2016; Wu et al., 2006). Manual tracings were developed to identify brain structures using neuroanatomical atlas and with the guidance of board certified neuroradiologists and neuroanatomists who were study investigators. Given the rigor used to identify structural boundaries, manual tracings are considered the “gold standard” for brain volumes (Bigler et al., 2010; Bigler, 2015) and are used as the comparison for volumes obtained using automated segmentation methods (Ambarki et al., 2011; Grimm et al., 2015; Lyden et al., 2016; Wu et al., 2006).

Importantly as a major limitation, manual tracing is time intensive to use, requires an extensive knowledge of neuroanatomy, software programs such as AFNI or ITK-SNAP, and requires extensive training with an expert rater (Wilke, de Haan, Juenger, & Karnath, 2011).
Reliability with an expert rater (e.g., neuroradiologist) is essential for reliable manual measurement of WMHs, and interrater reliability must be high with a Dice similarity coefficient > 0.85 (Dice, 1945; Theiss, Ridgewell, McHugo, Heckers, & Blackford, 2017). Given that manual tracing is so labor intensive, the use of automated methods that produce faster, reliable volumes has rapidly increased in recent years (Albright et al., 2016; Brewer, Magda, Airriess, & Smith, 2009; McCarthy et al., 2015; Wilke et al., 2011). To be able to make comparisons of MRI findings across imaging methods, the automated methods need to be comparable to the gold standard manually traced method (Fischl et al., 2002; Grimm et al., 2015; Schoemaker et al., 2016; Wilke et al., 2011).

**Automated Segmentation Methods**

Currently, automated segmentation tools are almost exclusively used in research settings, with limited incorporation into the clinical workflow (Brewer, 2009). Automated segmentation tools shorten the time it takes to obtain ROI volumes but often still require operator input and management for data formatting (Wilke et al., 2011). One benefit of automated methods is the potential to reduce rater error/bias due to the use of systematic segmentation algorithms, making them more appealing than clinical rating scales or manual tracing for replicability in research and clinical settings (Albright et al., 2016; McCarthy et al., 2015; Spitz et al., 2013).

**FreeSurfer.** Developed in 1999, FreeSurfer is an open-source automated program that segments and quantifies brain regions based on anatomic landmarks in structural imaging (Fischl, 2012; McCarthy et al., 2015). FreeSurfer’s fast and reliable procedure for obtaining ROI volumes was a welcome advancement in neuroimaging (Spitz et al., 2013). The FreeSurfer algorithm may over- or under-estimate the ROI, especially in brains with an abnormal size or shape, or major pathology (Schoemaker et al., 2016; Wenger et al., 2014). Each scan requires
significant time for preprocessing and segmentation (approximately 8 hours; Ochs et al., 2015; Wenger et al., 2014). FreeSurfer allows for visual inspection of automated tracings and simple boundary corrections to be made, which assist in increasing reliability of the results but require additional time and expertise (Fischl, 2012; McCarthy et al., 2015). While FreeSurfer is freely available to the public, FreeSurfer’s documentation states that it is not intended for clinical use, thereby limiting its use to research (https://surfer.nmr.mgh.harvard.edu; Ross et al., 2013).

**NeuroQuant®.** Developed by CorTechs Labs (www.cortechslabs.com) in 2009, NeuroQuant® is an outgrowth of FreeSurfer software (Ochs et al., 2015) and is the first image analysis program specifically designed for clinical use and has U.S. Food and Drug Administration approval (2006) (Albright et al., 2016). NeuroQuant® allows for fully automated quantification of brain volumes reducing the need for an onsite computer, technical support, and processing time (Brewer, 2009; Brewer et al., 2009). Reports are completed and returned to the requesting physician in approximately 5 to 10 minutes, or 20 minutes for the LesionQuant report (Albright et al., 2016; Azab et al., 2015; Ochs et al., 2015; Reid et al., 2017). NeuroQuant® has received Health Canada approval (licensed in 2016) and its coverage by Medicare in the United States, which serves as a guide for medical insurance coverage for other health providers (Brewer, 2009; Ross et al., 2013). With NeuroQuant®’s approval for clinical use, it has potential for widespread use including providing objective information about brain pathology in research and clinical settings (Albright et al., 2016).

The initial focus of NeuroQuant® was to characterize morphological changes in the hippocampi and temporal horns, due to their association with Alzheimer’s disease (Brewer et al., 2009; Kovacevic, Rafii, & Brewer, 2009). Given the numerous neurological and neuropsychiatric disorders where morphological differences occur, in 2015, CorTechs Labs
expanded NeuroQuant® to provide an in-depth volumetric analysis of MRI scans in individuals ages 3 to 100 years old (Luo, Airriess, & Albright, 2015). As of August 2016, CorTechs Labs also added the LesionQuant report, which assesses WMH volumes using T1 and FLAIR images (Luo et al., 2017a; Luo et al., 2017b). Originally designed to monitor WMH changes in multiple sclerosis, LesionQuant identifies WMHs via intensity differences using a threshold technique (Luo et al., 2017b). The LesionQuant report uses the same dynamic atlas as NeuroQuant® for preprocessing and normalization (Luo et al., 2017b) and shows good comparability with manual tracings in a multiple sclerosis sample (Luo et al., 2017b). The updated version of NeuroQuant® (NeuroQuant® 2.0) and the new LesionQuant report have not been evaluated in pediatric TBI. To avoid confusion, WMH volumes collected from the LesionQuant report will be referred to as simply NeuroQuant® WMH volumes in this study.

Dynamic atlas. Like FreeSurfer, NeuroQuant® uses an algorithm that includes: 1) a quality control step that checks that the uploaded images meet specifications required to perform automatic segmentation (i.e., 3-dimensional T1-weighted sequence, adequate contrast to differentiate ROIs); 2) biasfield and gradient non-linearity correction; and 3) automatic segmentation of ROIs using a probabilistic atlas (Kovacevic et al., 2009; Lyden et al., 2016). NeuroQuant® uses the same automated segmentation procedure as FreeSurfer; however, while FreeSurfer relies on a static atlas, NeuroQuant® uses a dynamic atlas for ROI segmentation (Ochs et al., 2015; White, Magda, Airriess, & Albright, 2015b).

FreeSurfer uses a static atlas (e.g., Deikan-Killiany atlas; https://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation), which is a single image that provides information about the expected spatial location, tissue class, and voxel intensity for any given ROI, and is created by averaging healthy control brains of a similar demographic (White,
Magda, Airriess, & Albright, 2015a). FreeSurfer’s default static atlas used during preprocessing consists of 39 adult brains (mean age of 38 years ± 10) for location and voxel intensity comparisons (Wenger et al., 2014).

NeuroQuant®’s dynamic atlas increases precision of ROI segmentation (see Figure 3) by creating a personalized atlas using healthy controls matched for age and sex of the individual, that summarizes voxel intensities, and anatomical locations (Albright et al., 2016; Ochs et al., 2015; White et al., 2015a). The healthy control sample consists of several thousand brains, both male and female, with an age range of 3 to 100 years which were obtained from publicly available neuroimaging databases, collaborative partners of CorTechs Labs, and other private datasets, to create a large normative database for NeuroQuant®’s dynamic atlas (White et al., 2015a).

*Figure 3.* Example of ROI segmentation using a static and dynamic atlas. Arrows are pointing to the thalamus and hippocampus. The dynamic atlas used age and sex matched controls for comparison to increase precision in ROI segmentations that are more age and sex appropriate ROI boundaries (see arrows). Used with permission by CorTechs Labs, Inc. (White et al., 2015b).
Comparison of Total WM and WMH Methods

Since WMHs are significantly associated with brain atrophy (Capizzano et al., 2004; Fujishima et al., 2014; Habes et al., 2016; Schmidt et al., 2005), and TBI is associated with reduced total WM volume (Königs et al., 2018) which contributes to brain atrophy (Cole et al., 2018; Königs et al., 2018), it is anticipated that total WM volume would be reduced with the presence of WMHs. If the presence of WMHs is associated with reduced WM volume, potentially, this means that total WM volumes captured a sufficient amount of WM pathology including what is associated with WMHs, and only total WM volumes would be sufficient to indicate WM integrity. This issue of potential redundancy of techniques needs to be assessed by comparing WMH volumes and total WM. It is unclear if the measure of WMH and WM are assessing the same construct or if they measure distinct things. That is do total WM volumes and WMHs assess two unique aspects of WM or do they assess the same WM construct using two different methods? Comparing the different methods used to assess WM becomes critically important in a clinical setting because of the financial cost for using a service like NeuroQuant®. The cost of all of the various NeuroQuant® reports if all performed on just one patient is approximately $450. In contrast, if only the report that includes WMH volumes are used, the cost would be only be $89 (www.cortechslabs.com).

Cognitive Function

Because of the variability of neural pathology, morbidities following pediatric TBI are heterogeneous (Allen et al., 2010; Smitherman et al., 2016). Chronic physical (Dahl & Emanuelson, 2013; Williams, Schache, & Morris, 2013), emotional (Max et al., 2012; NINDS, 2013) behavioral (Königs et al., 2018), and social deficits (Ryan et al., 2013; Wolfe et al., 2015; Yeates et al., 2013) are common following pediatric TBI. Cognitive impairments are among the
most disabling morbidities because of their direct effects on development, behavioral, social, and academic functioning, as well as potential for employment and independent living (Anderson et al., 2012; Tonks, Williams, Yates, & Slater, 2011). Cognitive impairments negatively affect educational attainment and financially independence in adulthood, resulting in financial burdens for the individual, their family and society, especially in individuals with severe injury (Arroyos-Jurado, Paulsen, Merrell, Lindgren, & Max, 2000; Catroppa et al., 2016; Lloyd et al., 2015). The Social Outcomes for Brain Injured Kids (SOBIK) study is a landmark study that assessed neuroimaging abnormalities and cognitive and social functions in pediatric TBI (Bigler et al., 2013a; Dennis et al., 2012; Shultz et al., 2016; Yeates et al., 2013). As the current study used the SOBIK dataset, further details of the SOBIK study and cognitive outcomes from that investigation are described below.

Normal cognition requires healthy neuronal activity from widely distributed brain regions and networks that work together in an integrated fashion (Ferrer et al., 2013; Filley & Fields, 2016; Kinnunen et al., 2011; Sharp, Scott, & Leech, 2014), and this communication occurs through the WM (Ferrer et al., 2013; Filley & Fields, 2016; Sharp et al., 2014). Given the importance of WM in cognitive functioning, it is clinically important to evaluate WM following TBI (Filley & Fields, 2016; Kinnunen et al., 2011; Sharp et al., 2014). While WM is related to all cognitive functions, the most straightforward and consistent relationship is likely with processing speed (Chevalier et al., 2015; Felmingham, Baguley, & Green, 2004; Prigatano, Gray, & Gale, 2008).

While the SOBIK investigation was focused on social outcome in TBI, select neuropsychological tests were obtained so as to characterize cognitive functioning in the sample by domains that included processing speed, intellectual functioning, attention, and language
COMPARISON OF WM METHODS IN PEDIATRIC TBI

(Bigler et al., 2013a; Dennis et al., 2012; Shultz et al., 2016; Yeates et al., 2013). The current investigation also explored how certain neuropsychological test findings related to the different methods for assessing total WM volume and WMHs. Since the current investigation capitalized on the existing SOBIK dataset, these analyses were constrained by what had already been acquired. A brief overview of the four domains of cognitive functioning, as assessed by SOBIK study, follows below.

**Processing Speed**

Processing speed is the speed at which new information can be processed, assessed, and a response formulated or recorded, that is how fast cognitive tasks can be completed (DeLuca & Kalmar, 2008; Gorman et al., 2016). Research suggests that processing speed is influenced by increased speed of signal transduction due to myelination (Chevalier et al., 2015) manifest by increased processing speed during development (Ferrer et al., 2013) and decline of processing speed in aging (Papp et al., 2014; Salthouse, 2000; van den Heuvel et al., 2006a). Processing speed can be affected by WM pathology in any brain region (Bigler et al., 2013a; Kochunov et al., 2010; Luo et al., 2016; Neitzel et al., 2016; Penke et al., 2010; Righart et al., 2013; Turken et al., 2008; van der Land et al., 2015). As such, slow processing speed is common following moderate to severe TBI where WM pathology is common (Felmingham et al., 2004; Prigatano et al., 2008); with increased injury severity resulting in greater disruptions of processing speed (Anderson et al., 2012; Prigatano et al., 2008).

There are multiple measures of processing speed. Processing speed is frequently assessed using timed tasks such as the Processing Speed Index (PSI) of the Wechsler Intelligence Scale for Children- Fourth Edition (WISC-IV; Wechsler, 2003b) which was used in the SOBIK study (Bigler et al., 2013a) and is the focus of the current investigation (additional details of the WISC-
IV PSI from the SOBIK are mentioned in the methods section). Processing speed deficits were associated with global brain atrophy in the SOBIK study (Bigler et al., 2013a); however, to date no study has specifically examined the relationship between total WM and WMHs and their association with processing speed.

Slow processing speed effects other cognitive deficits (Forn, Belenguer, Parcet-Ibars, & Ávila, 2008) by lengthening reaction time, disrupting the processing of new incoming information (Battistone, Woltz, & Clark, 2008; Gorman et al., 2016) and decreasing the amount of information processed within a given amount of time (Gorman et al., 2016; McAuley & White, 2011). The processing speed deficits adversely affect attention, decision-making, memory, thinking, social, and academic performance (Anderson, Catroppa, Morse, Haritou, & Rosenfeld, 2005) and the ability to perform complex cognitive manipulations (Chevalier et al., 2015; Fry & Hale, 1996, 2000). Children with processing speed deficits often lag behind their same aged peers in academic performance (Anderson et al., 2005). In addition to processing speed, other cognitive domains such as intellectual function, attention, and language may also be effected by WM pathology (Anderson et al., 2005; Anderson, Catroppa, Morse, Haritou, & Rosenfeld, 2009; Sullivan & Riccio, 2010).

**Intellectual Function**

The prevalence of intellectual impairments following TBI is a subject of debate, with some studies showing individuals with TBI having impaired intellectual function compared to controls (Araujo et al., 2017; Königs, Engenhorst, & Oosterlaan, 2016), others showing no significant difference between individuals with TBI and controls (Anderson et al., 2009). A third pattern can occur where intellectual impairments occur during the acute phase of injury, but recovers over time (Crowe, Catroppa, Babl, Rosenfeld, & Anderson, 2012). Data regarding the
association between WM and intellectual function is limited. This may be because intellectual function is often measured as a broad assessment of cognitive functioning (Ewing-Cobbs, Barnes, & Fletcher, 2003), which can be associated with varied pathology. Studies suggest that decreased intellectual function is associated with ventricular enlargement suggesting brain atrophy (Dennis et al., 2016; Johnson et al., 1994) and corpus callosum atrophy (Gale & Prigatano, 2010), while other studies find no association between intellectual function and white matter (Genc et al., 2017).

One of the most common measures of intellectual function in children is the Wechsler Abbreviated Scale of Intelligence (WASI; Genc et al., 2017; Strauss, Sherman, & Spreen, 2006; Wechsler, 1999). The SOBIK study used the WASI and found that participants with severe TBI had the lowest Full IQ scores (Shultz et al., 2016; Yeates et al., 2013), which were significantly lower than the IQ scores of the orthopedically injured control group (Bigler et al., 2015; Shultz et al., 2016). This supports previous research suggesting intellectual impairments in severe TBI (Anderson et al., 2009).

Intellectual impairments seem to be largely dependent on injury severity, with impairments in intellectual function being common following severe TBI (Anderson et al., 2009; Ewing-Cobbs et al., 2003). Children with mild TBI may have a better cognitive outcomes compared to adults, and worse outcomes following severe TBI (Anderson et al., 2009; Königs et al., 2016). Better outcomes following pediatric mild TBI may be due to neural plasticity and adaptive processes, whereas children with severe TBI often fail to meet developmental milestones, falling further behind their peers over time (Anderson, Morse, Catroppa, Haritou, & Rosenfeld, 2004; Ewing-Cobbs et al., 2003). A study of children with severe pediatric TBI found limited recovery of intellectual functioning 5 years post-injury which may be due to disruption of
normal brain development following the TBI (Anderson et al., 2009). Further, a meta-analysis of 81 studies suggest that children with severe TBI have worse intellectual function outcomes compared to adults (Königs et al., 2016). The heterogeneity of intellectual functioning across studies and injury severities makes prognosticating long-term intellectual outcome difficult.

**Attention**

Impaired attention can include difficulties in maintaining alertness (sustained attention), focusing or attending to target information (selective attention) and inhibition for irrelevant stimuli (Anderson et al., 2005; Posner & Petersen, 1990). Few studies have assessed the direct connection between WM pathology and attention. Disrupted neural networks in the frontal and temporal lobes (Anderson et al., 2005; Posner & Petersen, 1990), and in cerebral white matter (Araujo et al., 2017; Filley & Fields, 2016) have been associated with attentional deficits. Impaired attention can negatively affect performance on a wide range of neuropsychological and academic abilities (Anderson et al., 2005) by disrupting the encoding of incoming information, which can negatively affect both learning and memory (Allen et al., 2010; Sava, Paquet, Dumurgier, Hugon, & Chainay, 2016) Other cognitive impairments such as impaired memory and executive function as well as behavior problems and negative psychosocial outcomes are associated with poor attention (Anderson et al., 2005; Karver et al., 2012; Yeates et al., 2005).

Regardless of the association with WM, studies find impaired attention is a common consequence of pediatric TBI (Babikian & Asarnow, 2009; Ponsford & Kinsella, 1992), with increased injury associated with increased severity of impaired attention (Anderson et al., 2005; Catroppa & Anderson, 2005; Königs et al., 2018). In addition, the age of injury influences attentional outcomes, as younger children with TBI have more severe and persisting attentional deficits (Karver et al., 2012). The SOBIK study used the Test of Everyday Attention in Children
(TEA-Ch; Manly, Robertson, Anderson, & Nimmo-Smith, 1999) a widely used and valid measure in children with TBI (Anderson, Fenwick, Manly, & Robertson, 1998; Manly et al., 1999; Strauss et al., 2006) and found no difference in attention between the TBI and the orthopedically injured groups (Bigler et al., 2015).

**Language**

Successful communication and interaction is dependent on effective contextual comprehension and use of language (Rowley, Rogish, Alexander, & Riggs, 2017). Language deficits can occur in children and adults (Ewing-Cobbs et al., 2003; Lê, Coelho, Mozeiko, Krueger, & Grafman, 2012; Marini et al., 2011; Sullivan & Riccio, 2010). While some research has shown language to be affected by diffuse brain damage (Anderson & Yeates, 2010; Catroppa et al., 2016) there is more evidence that suggest language impairments are more location specific (Lê et al., 2012; Liégeois et al., 2013; Marini et al., 2011; Nagy et al., 2004; Walton, Dewey, & Lebel, 2018). Damage to the left hemisphere, and more specifically in the WM pathways connecting language regions (e.g., arcuate fasciculus) result in a variety of language impairments including impaired reading comprehension and word decoding (Barnes, Dennis, & Wilkinson, 1999; Liégeois et al., 2013). Additional areas such as the uncinate fasciculus and corpus callosum have also been areas of interest in the language literature, specifically in their relation to semantic deficits (Liégeois et al., 2013; Walton et al., 2018).

Children with TBI may develop language deficits (Lê et al., 2012; Marini et al., 2011). For example, research using a story retelling task found that children with TBI shared less essential information, omitted critical content, used fewer grammar elements, provided less structure to their story, had difficulty identifying the central idea, and had a reduced ability to organize their narrative discourse in a logical manner compared to participants without TBI.
COMPARISON OF WM METHODS IN PEDIATRIC TBI

(Chapman et al., 2004; Hay & Moran, 2005; Jorgensen & Togher, 2009; Lê et al., 2012). The Test of Language Competence-Expanded Edition was used in the SOBIK study to assess linguistic competence and detect language delays in children (Wiig & Secord, 1985). No publication to date has analyzed language scores from the SOBIK data.

Age at time of injury and injury severity may influence the severity of language deficits in children. Children who experience a TBI at a younger age have more significant language deficits than older children with TBI, which may be due to the disruption of normal language development (Sullivan & Riccio, 2010). Additionally, preschool aged children who sustained a severe TBI develop the most severe and persistent language deficits (Anderson et al., 2004; Liégeois et al., 2013; Posthuma, 2003; Sullivan & Riccio, 2010). The language deficits appear to be less severe if injury occurs at older ages, after the development of language skills has occurred (Sullivan & Riccio, 2010). In summary, injury severity, location of damage, and age at injury all predict language impairments following pediatric TBI (Lê et al., 2012; Liégeois et al., 2013; Marini et al., 2011).

Sex Differences

Studies of cognitive function indicate a consistent sex difference on certain cognitive tests. For example, healthy males tend to have higher scores than females on the PSI (processing speed) and the Creature Counting subtest of TEA-Ch (attention; Manly et al., 1999; Strauss et al., 2006). Given these known sex differences, it is important to assess sex differences in TBI populations, in order to fully understand the effects of brain injury and its relationship to WM.

Purpose

The purpose of this study is to assess WM integrity using the SOBIK study, to assess WMHs and total WM volumes following pediatric TBI based on new and existing clinical and
volumetric methods, WM methods have been validated for use in adult populations but have yet to be assessed for use in pediatric TBI. In order to determine which methods can be used in pediatric TBI, this study compared (1) a variety of WM assessment methods: including Scheltens rating scale, manual tracing, FreeSurfer, and NeuroQuant®; (2) assessed total WM volumes and WMHs to determine whether total WM volumes adequately capture WM pathology; and (3) assessed the relationship of WM ratings and volumes with select measures of cognitive function including processing speed, intellectual functioning, attention, and language in pediatric TBI participants.

**Aims and Hypotheses**

**Aim 1: Comparison of Quantitative White Matter Hyperintensity Methods**

If NeuroQuant® WMH volumes and manually-traced WMH volumes are robustly correlated they may equally provide volumetric information about WM integrity in pediatric TBI.

*Hypothesis 1.* Manually-traced WMH volumes will be positively correlated, and have good agreement with NeuroQuant® WMH volumes in pediatric TBI.

**Aim 2: Comparison of Qualitative and Quantitative White Matter Hyperintensity Methods**

No previous study involving pediatric TBI has examined the relationship between qualitative and quantitative measures of WM integrity; therefore, it is unclear how clinical ratings of WMHs relate to these quantitative methods. Since the clinical rating and quantitative method both assess WMH, it is hypothesized that the results of the two methods will be associated.

*Hypothesis 2:* Scheltens WMH ratings, will be positively correlated, and have good agreement with manually traced and NeuroQuant® WMH volumes in pediatric TBI.
Aim 3: Comparison of Quantitative Total White Matter Methods

Both NeuroQuant® and FreeSurfer use automated methods to obtain total WM volumes. FreeSurfer volumes have been compared to the gold-standard manually-traced WM volumes in adult (Fischl et al., 2002; Wenger et al., 2014) and pediatric samples (Bigler et al., 2010), with good agreement. NeuroQuant® has not.

Hypotheses 3. FreeSurfer total WM volumes will be positively correlated and have good agreement with NeuroQuant® total WM volumes in pediatric TBI.

Aim 4: Comparison of Total White Matter Volumes and White Matter Hyperintensities

In the basic multi-structural NeuroQuant® report (see Appendix A for example report), only total WM volumes but not WMH volumes are included. It is an added expense to obtain the FLAIR WMH volumes, which are included on a separate report (i.e., LesionQuant, see Appendix B for example report). This raises the question as to whether merely calculating total WM volume is sufficient to assess overall WM integrity.

Also, as other studies have demonstrated that WM volume loss is associated with injury severity (Königs et al., 2018; Narayana, 2017), it is important to know whether total WM volumes are significantly smaller in moderate-severe pediatric TBI compared to complicated-mild TBI and whether individuals who have WMHs have smaller total WM volumes. It was hypothesized there would be a significant difference in total WM volumes between injury severities, and between individuals with and without WMHs, where individuals with WMHs will have smaller total WM volumes.

Hypothesis 4a. Participants with moderate-severe TBI will have significantly smaller total WM volume compared to participants with complicated-mild TBI.
Hypothesis 4b. The subgroup with WMHs will have smaller total WM volumes compared to the subgroup without WMHs, in both complicated-mild and moderate-severe TBI groups.

**Research Question: Comparison of Cognitive Function and Methods of Assessing White Matter**

As discussed in the introduction, DAI underlies much of the WM pathology visible using MRI technology (Jarrett et al., 2016; Lee & Newberg, 2005; de la Plata et al., 2007) and can be quantified using automated methods to obtain total WM volumes, as well as clinical and quantitative methods for assessing WMHs. While the above four hypotheses outline the primary objectives and goals, the SOBIK participants also completed a neuropsychological battery including measures of processing speed, intellectual function, attention, and language (Bigler et al., 2013a; Heverly-Fitt et al., 2014; Shultz et al., 2016; Yeates et al., 2013). As such, a research question was designed to assess the relationships between WM methods and cognitive functions.

Because of the multiple assessments of WM integrity and cognitive outcomes to potentially explore, the SOBIK sample size was too limited for hypotheses with multiple comparisons. Nonetheless, a descriptive analysis was appropriate and provided information within the SOBIK cohort regarding associations between total WM volumes and WMHs and cognitive test scores. As such, this research question that was not hypothesis driven but rather descriptive in nature as to how these methods of assessing white matter integrity (Scheltens rating scale, manual tracing, NeuroQuant®, and FreeSurfer) relate to measures of cognitive functioning (measures of processing speed, intellectual functioning, attention, and language). No prior studies in pediatric TBI have explored multiple methods for assessing WM integrity and
cognitive outcomes in the chronic injury phases. As such, this aspect of the dissertation also provides novel information.

As individual comparisons of each WM metric would dramatically increase the risk of false positive correlations (or type 1 error), the current analyses employed a correlational SEM approach to simultaneously assess WM methods and cognitive functions. While this descriptive research question has no hypothesis, based on the literature processing speed probably has the most consistent observed relationship with WM, regardless of etiology (Genc et al., 2017; Jacobs et al., 2013; Kuznetsova et al., 2016). As such, it is anticipated that processing speed may have the strongest relationship with WM metrics derived from the SOBIK study using this SEM approach. From the results, only the strongest relationships between WM methods and cognitive measures were explored.

Additionally, as previous research has shown that there are potential sex differences on the PSI (processing speed) and the Creature Counting subtest of TEA-Ch (attention; Manly et al., 1999; Strauss et al., 2006), sex differences in the relationship between WM metrics and cognitive functions were also assessed using a multi group analysis of the same SEM model.

Methods

Participants

This study used archival data collected prospectively from 2006 to 2010 for the SOBIK. The SOBIK investigation had IRB approval from each participating institution, including Brigham Young University (Bigler et al., 2013a; Bigler et al., 2016; Dennis et al., 2012; Heverly-Fitt et al., 2014; Shultz et al., 2016; Yeates et al., 2013). Each participant’s parent or legal guardian provided written informed consent. The current study also has Brigham Young University IRB approval.
SOBIK inclusion and exclusion criteria. The SOBIK study enrolled children who had a TBI with a GCS of ≤12 on the day of injury, or a GCS scores of 13-15 with confirmed brain insult or skull fracture (Bigler et al., 2013a; Bigler et al., 2016; Dennis et al., 2012; Shultz et al., 2016; Yeates et al., 2013). The SOBIK study inclusion criteria included children who had a TBI between 8 to 13 years of age, admitted to hospital for at least 24 hours, and the TBI did not occur prior to age 4 and who were enrolled in the study one to four years after their injury. The SOBIK study exclusion criteria included placement in special education, homeschooled, prior TBI, prior neurological disorders, language disability, mental retardation, history of child abuse or assault, psychological disorders requiring hospitalization, sensory or motor impairment, non-English speaking, or met MRI exclusion criteria (Bigler et al., 2013a; Bigler et al., 2013b; Dennis et al., 2012; Shultz et al., 2016; Yeates et al., 2013). The neuropsychological test session and MRI acquisition occurred on the same day.

Current study inclusion and exclusion criteria. For the current study, additional inclusion criteria included a GCS score at the time of injury, usable neuroimaging data (i.e., T1, FLAIR scans) and available neuropsychological test scores. Participants were excluded if their scans had significant motion or scanner artifact, or incomplete imaging sequences. Participants with unusable FLAIR images were still included if they had usable T1 images.

Demographic and medical data. Demographic data included age at injury and neuropsychological testing, maternal education (years), ethnicity, and sex. Medical data included GSC score (severe: score of 8 or less, moderate: scores from 9 to 12, complicated mild: score from 13 to 15 with CT confirmed brain pathology), hospital length of stay, and imaging data.
Neuropsychological Tests

While cognitive functions are frequently impaired following pediatric TBI (Anderson, Catroppa, Rosenfeld, Haritou, & Morse, 2000; Bigler et al., 2015; Gorman et al., 2016; Keightley et al., 2014a), this study focused only on the cognitive domains assessed in the SOBIK study which are: intellectual function, attention, processing speed, and language. Intellectual function was assessed using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), attention was assessed using the Test of Everyday Attention for Children, (Manly et al., 1999), processing speed was assessed using the Wechsler Intelligence Scale for Children, 4th Edition, Processing Speed Index (WISC-IV, PSI; Wechsler, 2003a), and language was assessed using the Test of Language and Competence- Expanded Edition (TLC-E; Wiig & Secord, 1985). Table 2 describes the neuropsychological tests administered in the SOBIK study (Bigler et al., 2013a; Heverly-Fitt et al., 2014; Shultz et al., 2016). See Appendix C for detailed description of each tests. Because the TEA-Ch does not have a composite score and the three subtests are significantly correlated with each other (Manly et al., 1999; Strauss et al., 2006), the scaled scores for the three subtests were summed and averaged, creating an overall attention score which had a mean of 10 and SD of 3 (Bigler et al., 2015). The overall attention score was included in the data analysis.
Table 2

Complete List of Cognitive Tests Obtained in the SOBIK dataset and Used in the Current Study

<table>
<thead>
<tr>
<th>Test</th>
<th>Subtests Administered</th>
<th>Cognitive Function Assessed</th>
<th>Composite Score</th>
<th>Scores M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wechsler Abbreviated Scale of Intelligence (WASI)</td>
<td>Vocabulary</td>
<td>Intelligence</td>
<td>Full Scale Intelligence Quotient (FSIQ)</td>
<td>100 (15) *</td>
</tr>
<tr>
<td>(Wechsler, 1999)</td>
<td>Matrix Reasoning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test of Everyday Attention for Children (TEA-Ch)</td>
<td>Creature Counting†</td>
<td>Attention</td>
<td>Subtests scores were averaged</td>
<td>10 (3)**</td>
</tr>
<tr>
<td>(Manly et al., 1999)</td>
<td>Walk Don’t Walk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code Transmission</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wechsler Intelligence Scale for Children, 4th Edition, Processing Speed Index (WISC-IV, PSI) (Wechsler, 2003a)</td>
<td>Symbol Search</td>
<td>Processing speed</td>
<td>Processing Speed Index†</td>
<td>100 (15)*</td>
</tr>
<tr>
<td>(Wiig &amp; Secord, 1985)</td>
<td>Cancellation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Wiig &amp; Secord, 1985)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * = standardized scores below 85 are considered impaired; ** = scaled scores below 7 are considered impaired. † = sex difference noted with males perform better than females.
Magnetic Resonance Imaging

All participants underwent neuroimaging a minimum of 1 year post injury (average 2.6 years post injury; range 1 to 5.2 years) using either a 1.5 T GE Signa Excite scanner or 1.5 T Siemens Symphony MRI scanner. Each site used comparable imaging sequences to obtain a dual-echo proton density T2-weighted (T2), FLAIR, and T1-weighted images (T1; Bigler et al., 2013a; Yeates et al., 2013; see Figure 4). Total WM volumes were obtained using the T1 (TR = 3000 ms; TE = 4.0 ms; flip angle = 8; 170 slices at 1.2 mm thick; field of view [FOV] = 24; matrix = 256 x 256 x 124), WMH volumes were obtained using the FLAIR (TR = 3000 ms; TE = 100 ms; flip angle = 90; 24 slices at 5 mm thick; FOV = 24; matrix = 256 x 256). The T2 scans (TR = 3000; TE1 = 12 ms; TE2 = 100 ms; flip angle = 90; 3 mm interleaved, FOV = 24; matrix = 256 x 256), were used to confirm the presence of WMHs on FLAIR images during manual tracing. Figure 4 shows WM neuroimaging methods. To ensure quality data acquisition, phantom imaging was done at each neuroimaging site (Bigler et al., 2013a; Yeates et al., 2013).
COMPARISON OF WM METHODS IN PEDIATRIC TBI

WMH Methods

- **Scheltens**
  - Requires: FLAIR
  - WMH Ratings

- **Manual Tracing**
  - Requires: FLAIR, T2, T1
  - Preprocessed FLAIR, T2

- **NeuroQuant®**
  - Requires: FLAIR, T1
  - Dynamic Atlas

LesionQuant Report

Total WM Methods

- **FreeSurfer**
  - Requires: T1
  - Recon-all

- **NeuroQuant®**
  - Requires: T1
  - Dynamic Atlas

Multi-Structural Report
**Figure 4.** Diagram summarizing image sequences used for each WM neuroimaging methods. Scheltens used raw FLAIR images for visual ratings. Manual tracing used FLAIR, T2, and T1 images. The preprocessed FLAIR images were used for manual tracings, and the preprocessed T2 was used to confirm the presence of WMHs for manual tracings. NeuroQuant® WMH method required the raw FLAIR and T1 images. * = hyperintensity in right lateral frontal lobe is an old contusion previously classified as gray matter hyperintensity, † = image does not align with remaining images as the provided image from the report was unable to be manipulated.

**Preprocessing of MR images.** Manual and automatic methods have slightly differing approaches in preprocessing the raw DICOM images. Scheltens ratings used raw DICOM images and required no preprocessing. Both FreeSurfer and NeuroQuant® used automatic within method preprocessing, whereas manual tracing required operator input.

**Manual tracing.** The raw DICOM images from the SOBIK dataset were preprocessed using Advanced Normalization Tools (ANTs) toolset. The raw DICOM images were converted to NiFti files and then T1 images were AC-PC aligned, biasfield corrected, and resampled to 1 mm isotropic value. The FLAIR and T2 images were affine registered to the preprocessed T1 image.

**NeuroQuant®.** Images were processed by NeuroQuant® using a dynamic atlas, which takes into account the participant’s sex and age, and creates a custom target image specific to the individual for optimal segmentation and comparison (White et al., 2015a).

**FreeSurfer.** The raw DICOM images were converted to NiFti files. Using the T1 image, FreeSurfer automatically strips away the skull, corrects for motion and artifacts, and carries out gray-WM segmentation followed by ROI segmentation (Fischl, 2012; McCarthy et al., 2015).
For a detailed list of processing procedures by recon-all in FreeSurfer, see https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all.

**Qualitative Rating of White Matter Hyperintensities**

**Scheltens rating scale.** The Scheltens rating scale was used to assess hyperintensities (see Appendix D for example of full Scheltens rating scale; Scheltens et al., 1993). The Scheltens rating scale was modified to exclude ratings of gray matter areas. The modified rating scale scores ranged from 0 to 36, with 36 indicating extensive WMHs (see Table 3 for the modified rating scale). Two raters independently rated all brain images and had a high interrater reliability (Dice Similarity Coefficient >.9; see Appendix E for example ratings). The Scheltens ratings yield semi-quantitative ratio data. The Scheltens ratings for the SOBIK dataset were previously collected but had not previously analyzed or published. A SOBIK investigator (EDB) reviewed all scans and Scheltens ratings prior to inclusion of the Scheltens data in the current study.
Table 3

*Modified Scheltens Rating Scale for White Matter Hyperintensities (Scheltens et al., 1993)*

<table>
<thead>
<tr>
<th>Region</th>
<th>Scale</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periventricular hyperintensities (PVH 0-6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps occipital</td>
<td>0 to 2</td>
<td>0 = absent</td>
</tr>
<tr>
<td>frontal</td>
<td>0 to 2</td>
<td>1 = ≤ 5 mm</td>
</tr>
<tr>
<td>Bands lateral ventricles</td>
<td>0 to 2</td>
<td>2 = &gt; 5 mm and &lt; 10 mm</td>
</tr>
<tr>
<td>White matter hyperintensities (WMH 0-24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>0 to 6</td>
<td>0 = absent</td>
</tr>
<tr>
<td>Parietal</td>
<td>0 to 6</td>
<td>1 = &lt; 3 mm, n ≤ 5</td>
</tr>
<tr>
<td>Occipital</td>
<td>0 to 6</td>
<td>2 = &lt; 3 mm, n &gt; 6</td>
</tr>
<tr>
<td>Temporal</td>
<td>0 to 6</td>
<td>3 = 4-10 mm, n ≤ 5</td>
</tr>
<tr>
<td>Internal capsule</td>
<td>0 to 6</td>
<td>4 = 4 mm, n &gt; 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 = &gt; 11 mm, n &gt; 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 = confluent</td>
</tr>
<tr>
<td>Total score</td>
<td>0 to 36</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Semi-quantitative rating of signal hyperintensities by brain region, with the range of the scale between brackets; n = number of lesions.

**Volume Acquisition**

*Manual tracing.* While manual tracing is the gold standard for volumetric analyses, a gold standard method or protocol for identifying WMHs currently does not exist (Schmidt et al., 2012). This is likely because WMHs vary tremendously in size, location, and shape. Other studies that have obtained WMH volumes using manual tracings simply state that WMHs were traced by expert raters who were trained by and achieved agreement with experts (e.g. neuroradiologists, neurologists; Capizzano et al., 2004; Vannorsdall et al., 2009). This study also
collected WMHs volumes manually using expert raters. Using AFNI, WMHs were manually traced, slice by slice on the affine registered FLAIR images (see Figure 4). Two investigators, blind to injury severity, independently traced WMH on all brain images and achieved a high inter-rater reliability of >.9 Dice similarity coefficient. Manual tracings followed the WMH boundaries and were visually inspected to ensure uniformity. Periventricular regions were not included in the WMH volumes, unless they were asymmetrical and extended into deep WM. Discrepancies between investigators were adjudicated by a neuroimaging expert (EDB). Scans with discrepancies were retraced until an agreement on WMH boundaries were reached. The total WMH voxel volumes (mm3) were converted to cm3 to allow for between method comparisons.

**NeuroQuant®.** NeuroQuant® provides a number of structural reports that assess specific volumetric changes compared to their normative database (normal age and sex dynamic atlas). For this study, the Multi Structure Atrophy and LesionQuant reports were used. The Multi Structure Atrophy report provides total WM volumes using the T1 image (see Appendix A for example report). The LesionQuant report provides WMH volumes using both T1 and FLAIR images (see Appendix B for example report). LesionQuant uses a thresholding algorithm to identify WMHs (Luo et al., 2017b).

**FreeSurfer.** Following the method details from (Bigler et al., 2010) for volumetric segmentation, FreeSurfer v6.0.0 (surfer.nmr.mgh.harvard.edu) was used to collect total WM volume and total intracranial volumes (TICV) for head size correction. Results were visually inspected to ensure the accuracy of the skull strip, registration, and WM segmentation. As manual edits can be used to optimize accuracy (McCarthy et al., 2015), volumes following manual edits were also collected, in addition to unedited volumes. Using the toolbox in
Freeview, manual edits consisted of 1) adding back in portions of the brain that were excluded as part of the skull strip, and 2) adding or subtracting WM voxels by using “Recon Edit” paint tool (Bigler et al., 2010; McCarthy et al., 2015). Following editing, images were reprocessed and visually inspected for accuracy before edited volumes were collected.

**Total intracranial volume correction.** To account for variability in head size WMH and total WM volumes were divided by TICV and multiplied by 100 to provide whole number values (Tate, Khedraki, Neeley, Ryser, & Bigler, 2011). NeuroQuant® volumes are automatically corrected in this same manner, and FreeSurfer automatically collect TICV within method. Manual tracing of the TICV were obtained following the methods outlined by Eritaia et al. (2000).

**Statistical Analysis**

Descriptive statistics data were reported as mean ± SD or percent as appropriate. Analyses were performed using Statistical Package for the Social Sciences (SPSS) 24.0.0.0, and Microsoft® Excel for Mac 15.23.2 on an Apple Macintosh computer with OS X version 10.10.5 (“Yosemite”). Correlations and intra-class correlations (ICC) were performed using SPSS and SEM models were performed using Analysis of Moment Structures (AMOS), a subprogram of SPSS. Bland-Altman plots were created using the batplot command program in STATA 14.2. A PC computer with Windows10, version 1703 was used for AMOS and STATA analyses.

**Missing Data**

Four participants were missing or had MRI scans that were not usable (2 complicated-mild, 2 severe) and 4 participants had missing neuropsychological test scores (1 complicated-mild, 2 moderate, 1 severe). There was no overlap in participants missing MRI or neuropsychological test data. The missing MRI data was due to one participant who did not have
FLAIR imaging, substantial artifact in 2 participants, and one participant in which the automated algorithms were unable to process the images despite multiple attempts, likely due to significant brain pathology (Schoemaker et al., 2016; Sidaros et al., 2009; Wenger et al., 2014). Missing imaging data was not imputed; however, the maximum likelihood estimate was used to account for missing data in the SEM model comparing imaging methods to neuropsychological data.

**Statistical Analyses for Normality**

Based on the Kolmogorov-Smirnov test of normality, WMH volumes and ratings were non-normally distributed: Scheltens $D(56) = .20, p = .00$; manual tracing $D(56) = .40, p = .00$; NeuroQuant® $D(55) = .20, p = .00$. Non-normal distributions were adjusted as required in each statistical analysis using log transformations (Bland & Altman, 1986; Bland & Altman, 1995; Bland & Altman, 2010; Giavarina, 2015; Rankin & Stokes, 1998). The log transformation resulted normal distributions for manual tracing ($D[40] = .14, p = .054$) and NeuroQuant® ($D[58] = .11, p = .09$), but not for Scheltens ratings ($D[50] = .15, p = .006$).

**Comparison of Quantitative White Matter Hyperintensity Methods**

To assess the reliability between methods, both correlation and agreement were assessed (Koo & Li, 2016; Rankin & Stokes, 1998; Wenger et al., 2014). Correlations were used to assess the relationship (whether there is and the strength of the relationship) the between two variables, whereas agreement based on the assumption that the two data sets have a commonality with a common variance and mean (Giavarina, 2015; Wenger et al., 2014). Results were considered significant at $p \leq .05$.

It was hypothesized that manual traced WMH volumes would be positively correlated, and have good agreement with NeuroQuant® WMH volumes in pediatric TBI (see Figure 5). Due to the non-normally distributed data, a Kendall’s tau correlation was used to assess the
association between manually traced and NeuroQuant® WMH volumes. To assess absolute agreement between WMH volumes a two-way mixed ICC was used (Cicchetti, 1994; Shrout & Fleiss, 1979; Wenger et al., 2014). Absolute agreement is considered poor when the ICC is below .40, fair between .40 and .59, good between .60 and .74, or excellent between .75 and 1.00 (Cicchetti & Sparrow, 1981). A Bland-Altman plot with a 95% limited agreement was also used to visually assess and describe the agreement between NeuroQuant® and gold standard manual tracing (Bland & Altman, 1986). A good agreement is a mean difference close to zero and the majority of the data lying within 2 SD of the mean difference or 95% limits of agreement (Bland & Altman, 1986; Bland & Altman, 1995; Bland & Altman, 2010; Giavarina, 2015). To correct for the non-normal data, data for the ICC and Bland-Altman plot were log transformed as recommended by several studies (Bland & Altman, 1986; Bland & Altman, 1995; Bland & Altman, 2010; Euser, Dekker, & Cessie, 2008; Giavarina, 2015). Because log transformations exclude values of 0 which represent participants identified as having no WMHs, analyses of agreement only included participants who had WMHs.

Figure 5. A visual of hypotheses in Aims 1 and 2.
Comparison of Qualitative and Quantitative White Matter Hyperintensity Methods

It was hypothesized that WMH ratings using the Scheltens rating scale, would be positively correlated, and have good agreement with manually traced and NeuroQuant® WMH volumes in pediatric TBI (see Figure 5). All imaging data were converted to z-scores due to the varied metrics (i.e., ratings, volumes) between methods (Königs et al., 2018). Due to the non-normal distribution of the data a Kendall’s tau correlation was used to assess the relationship between manually traced and NeuroQuant® WMH volumes with Scheltens rating of WMH. Absolute agreement between z-scores were determined by ICC (Cicchetti, 1994; Shrout & Fleiss, 1979; Wenger et al., 2014) and Bland-Altman plot with 95% limited agreement (Bland & Altman, 1986). To meet the assumptions of normality, the data for the ICC and Bland-Altman plots were log transformed.

Comparison of Total White Matter Volume Methods

It was hypothesized that FreeSurfer total WM volumes would be positively and significantly correlated and have good agreement with NeuroQuant® total WM volumes in pediatric TBI (see Figure 6). As NeuroQuant® and FreeSurfer total WM volumes were normally distributed, a Pearson’s R was used to assess the linear correlation between NeuroQuant® and FreeSurfer total WM volumes, and an ICC and Bland-Altman plot was used to determine absolute agreement (see Figure 6; Bland & Altman, 1986; Cicchetti, 1994; Shrout & Fleiss, 1979; Wenger et al., 2014). The log transformed data was used for the Bland-Altman plot as the mean difference was non-normally distributed (Bland & Altman, 1986; Bland & Altman, 1995; Bland & Altman, 2010; Giavarina, 2015; Rankin & Stokes, 1998).
Comparison of Total White Matter Volumes by Injury Severity

It was hypothesized that participants with complicated-mild TBI will have significantly larger total WM volume compared to participants with moderate-severe TBI (see Figure 7). An independent samples t-test was used to compare total WM volume in complicated-mild with moderate-severe TBI participants using injury severity as the dependent variable and total WM volumes as the independent variable. Results were considered significant at $p \leq .05$. A Cohen’s $d$ was used to measure the effect size.

Comparison of Total White Matter Volumes and White Matter Hyperintensities by Injury Severity

It was hypothesized that participants without WMHs will have larger total WM volumes compared to participants with WMHs, in both complicated-mild and moderate-severe TBI groups (see Figure 8). Independent samples t-tests were used to compare total WM volumes of
participants with and without WMHs, in both the complicated-mild and moderate-severe TBI groups. The dependent variables are the presence of WMH and the independent variables are the total WM volumes. Results were considered significant at $p \leq .05$. A Cohen’s d was used to measure the effect size.

**Figure 8.** Visual of hypotheses 4b: Comparing total WM volumes in those with and without WMHs in both complicated-mild and moderate-severe TBI participants; WMH (-) = WMH not present; WMH (+) = WMH present.

**Research Question: Comparison of Cognitive Function and Methods of Assessing White Matter**

As a descriptive part of the dissertation, an SEM model was used to assess the associations between WMH ratings (Scheltens), WMH volumes (manual and NeuroQuant®), and total WM volumes (NeuroQuant® and FreeSurfer) with scores on measures of intelligence, attention, processing speed, and language (see Figure 9). All imaging data were converted to z-scores due to differing metrics across cognitive tests and WM methods (ratings vs. volumes). A Benjamini Hochberg test was used to correct for Type 1 error, which has a false discovery rate (FDR) of .05 (Benjamini & Hochberg, 1995). Given that sex differences for the attention and
processing speed measures are well described (Manly et al., 1999; Strauss et al., 2006), the chi-squared difference test from a multiple group analysis SEM model was used to assess the associations between WM methods and cognitive functions accounting for sex.

*Figure 9.* Visual of the research question: All variables are co-varied. The multiple group analysis assessed sex differences and the associations between WM methods and cognitive functions.
Results

Demographic Data

Sixty children with TBI had usable brain MRI scans meeting all inclusion and quality assurance criteria for this study. Of the 60 TBI participants, one participant had only T1 images. Table 4 shows the demographic data for study participants. The mean age at the time of injury was 7.9 years and the mean age at the time of neuroimaging and neuropsychological testing was 10.5 years. Thirty-two (53%) participants had complicated-mild TBI, 9 (15%) participants had moderate TBI, and 19 (32%) participants had severe TBI. There was no difference between complicated-mild and moderate-severe TBI groups for sex ($p = .22$), maternal education ($p = .14$), age at injury ($p = .75$), age at testing ($p = .77$), and time since injury to testing ($p = .70$). Not surprisingly, hospital length of stay was significantly different ($p = .001$), as individuals with moderate-severe TBI were more likely to have longer lengths of stay compared to individuals with complicated-mild TBI.
Table 4

Demographic and Medical Data of TBI Children

<table>
<thead>
<tr>
<th>Demographics (N = 60)</th>
<th>M (SD) or N%</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at testing (years)</td>
<td>10.5 (1.5)</td>
<td>8.1 – 13.2</td>
</tr>
<tr>
<td>Age at injury (years)</td>
<td>7.9 (1.9)</td>
<td>3.2 – 12.1</td>
</tr>
<tr>
<td>Time since injury (years)</td>
<td>2.6 (1.2)</td>
<td>1.0 – 5.2</td>
</tr>
<tr>
<td>Hospital length of stay (days)</td>
<td>5.5 (9.3)*</td>
<td>1 – 58</td>
</tr>
<tr>
<td>Maternal education (years)</td>
<td>13.5 (2.3)</td>
<td>8 – 18</td>
</tr>
<tr>
<td>Sex (% Male)</td>
<td>65%</td>
<td></td>
</tr>
</tbody>
</table>

**Ethnicity**

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>N%</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>78%</td>
</tr>
<tr>
<td>Black</td>
<td>7%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>7%</td>
</tr>
<tr>
<td>Multiracial/nonspecific</td>
<td>8%</td>
</tr>
</tbody>
</table>

**GCS Score**

<table>
<thead>
<tr>
<th>GCS Score</th>
<th>M (SD)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complicated-Mild (n = 32)</td>
<td>14.8 (0.5)</td>
<td>13 – 15</td>
</tr>
<tr>
<td>Moderate-Severe (n = 28)</td>
<td>5.8 (3.4)</td>
<td>3 – 12</td>
</tr>
</tbody>
</table>

**MRI Imaging**

<table>
<thead>
<tr>
<th>MRI Imaging</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Both T1 &amp; FLAIR (n = 59)</td>
<td></td>
</tr>
<tr>
<td>Only T1 (n = 1)</td>
<td></td>
</tr>
</tbody>
</table>

*Note. TBI = traumatic brain injury; GCS = Glasgow Coma Scale, * = significant difference between complicate-mild and moderate-severe TBI.*

**Ratings and Volumetric Descriptive Data**

WM integrity was assessed using five different imaging methods. Descriptively, Table 5 shows WMH ratings and volumetric data by injury severity and total group. Participants with moderate-severe TBI had significantly higher Scheltens ratings (p = .01) and smaller FreeSurfer
total WM volumes \((p = .01)\) than participants with complicated-mild TBI (see Table 5). There was no difference between the edited FreeSurfer total WM volumes and the raw, un-edited FreeSurfer volumes \((t = -.23, df = 58, p = .82)\). As such, the raw, un-edited FreeSurfer volumes were used for data analysis.

Table 5

*Ratings or Volumetric Data Adjusted for TICV within Method, Grouped by Injury Severity and Total Group*

<table>
<thead>
<tr>
<th>Method</th>
<th>Complicated-Mild ((n = 32))</th>
<th>Moderate-Severe ((n = 28))</th>
<th>Total Group ((N = 60))</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMH present**</td>
<td>63%</td>
<td>74%</td>
<td>70%</td>
</tr>
<tr>
<td>Scheltens WMH Ratings</td>
<td>1.90 (1.81)</td>
<td>5.25 (6.13)*</td>
<td>3.49 (4.69)</td>
</tr>
<tr>
<td>Manual Tracing- WMH Volumes (cm³)</td>
<td>.01 (.01)</td>
<td>.07 (.25)</td>
<td>.04 (.17)</td>
</tr>
<tr>
<td>NeuroQuant®- WMH Volumes (cm³)</td>
<td>.07 (.07)</td>
<td>.08 (.09)</td>
<td>.07 (.08)</td>
</tr>
<tr>
<td>NeuroQuant®- Total WM Volumes (cm³)</td>
<td>27.83 (2.42)</td>
<td>28.40 (2.49)</td>
<td>28.10 (2.45)</td>
</tr>
<tr>
<td>FreeSurfer- Total WM Volumes (cm³)</td>
<td>29.04 (1.43)</td>
<td>27.89 (1.84)*</td>
<td>28.51 (1.72)</td>
</tr>
</tbody>
</table>

*Notes* Volume (cm³) are adjusted for total intracranial volumes (TICV); WMH = white matter hyperintensities; WM = white matter; volumes are cm³, * = significantly different from the complicated-mild group; ** = WMHs identified using manual tracings.

**Neuropsychological Descriptive Data**

Neuropsychological descriptive data are shown by injury severity (see Table 6).

Interestingly, participants with moderate-severe TBI had better scores for intellectual function,
processing speed, and language compared to participants with complicated-mild TBI. There was a significant difference between injury severities for attention ($p = .05$). For all TBI participants, cognitive impairment (for descriptive purposes will be defined as scores 2 SD below the mean) occurred in 36% for attention, 34% for language, 17% for intellectual function, and 12% for processing speed.

Table 6

*Neuropsychological Data Grouped by Injury Severity and Total Group*

<table>
<thead>
<tr>
<th>Neuropsychological Test</th>
<th>Complicated-Mild ($n = 32$)</th>
<th>Moderate-Severe ($n = 28$)</th>
<th>Total Group ($N = 60$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wechsler Abbreviated Scale of Intelligence (WASI) FSIQ</td>
<td>98.91 (15.10)</td>
<td>100 (13.04)</td>
<td>99.41 (14.08)</td>
</tr>
<tr>
<td>Test of Everyday Attention for Children (TEA-Ch)</td>
<td>8.63 (2.85)</td>
<td>7.18 (2.62)*</td>
<td>7.94 (2.81)</td>
</tr>
<tr>
<td>Wechsler Intelligence Scale for Children, 4th Edition – Processing Speed Index (WISC-IV PSI)</td>
<td>100.88 (12.11)</td>
<td>101.12 (14.37)</td>
<td>100.98 (13.05)</td>
</tr>
<tr>
<td>Test of Language and Competence- Expanded Edition (TLC-E)</td>
<td>7.68 (2.36)</td>
<td>8.08 (2.83)</td>
<td>7.86 (2.56)</td>
</tr>
</tbody>
</table>

*Note.* WASI and the WISC-IV PSI are standardized scores and the TEA-Ch and TLC-E are scaled scores; * = significantly different from the complicated-mild group.
Comparison of Quantitative White Matter Hyperintensity Methods

A poor agreement was found between NeuroQuant® and manual tracing (ICC [39] = .24, CI = -.19 to .56). The Bland-Altman plot mean difference was .79 with 2/39 data points outside the 95% limit of agreement (see Figure 10). The Kendall’s tau correlation comparing manually traced and NeuroQuant® WMH volumes was not significant (r = .12, p = .21, n = 56; see Figure 10). Similar results were found when outliers were excluded (see Appendix G).

<table>
<thead>
<tr>
<th>Scatter Plots</th>
<th>Bland-Altman Plots</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>NeuroQuant vs. Manual WMH Volumes (cm³)</td>
<td>NeuroQuant vs. Manual Tracing (log)</td>
</tr>
<tr>
<td>Manual Tracings</td>
<td>Mean</td>
</tr>
<tr>
<td>R² = .33</td>
<td>Difference</td>
</tr>
<tr>
<td>B</td>
<td></td>
</tr>
<tr>
<td>NeuroQuant WMH volumes (cm³) vs. Scheltens Ratings</td>
<td>NeuroQuant vs. Scheltens (z-log)</td>
</tr>
<tr>
<td>Scheltens Ratings</td>
<td>Mean</td>
</tr>
<tr>
<td>R² = .46</td>
<td>Difference</td>
</tr>
</tbody>
</table>
Figure 10. For WM method comparisons, correlational scatterplots are to the left and Bland-Altman plots assessing agreement are to the right. The scatterplots include a ‘line of best fit’ (dotted line) for the correlation. The Bland-Altman plots show upper and lower limit (dotted lines) defined as 1.96 standard deviation and the mean difference (black line). The mean
difference is close to zero for all comparisons, with the farthest difference for the comparison between NeuroQuant® and manual tracings.

**Comparison of Qualitative and Quantitative White Matter Hyperintensity Methods**

Excellent agreements were found between both manually traced and Scheltens rating z-scores ($ICC [57] = .91, CI = .85 to .95$) and NeuroQuant® and Scheltens z-scores ($ICC [56] = .79, CI = .64 to .88$). The Bland-Altman plot comparing manually traced WMH volumes and Scheltens ratings mean difference was -.15 with 3/37 data points outside the 95% limit of agreement of -1.59 and 1.30. The Bland-Altman plot comparing NeuroQuant® WMH volumes and Scheltens ratings mean difference was .06 with 2/49 data points outside the 95% limit of agreement of -1.89 and 2.02 (see Figure 10 for Bland-Altman plots). Kendall’s tau correlations comparing manually traced WMH volumes and Scheltens WMH ratings were significant ($r = .50, p = .00, n = 57$), as were correlations comparing NeuroQuant® and Scheltens WMH ratings ($r = .29, p = .005, n = 56$; see Figure 10). Similar results were found when outliers were excluded (see Appendix G). For ease of comparison, a qualitative summary of the correlation and agreement findings of the WMH method comparisons for the first two hypotheses is provided in Table 7.
Table 7

Review of WMH Method Comparison Findings

<table>
<thead>
<tr>
<th>Methods</th>
<th>Scheltens Ratings</th>
<th>NeuroQuant® WMH</th>
<th>Manual Tracing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation</td>
<td>Agreement</td>
<td>Correlation</td>
</tr>
<tr>
<td>Scheltens Ratings</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>NeuroQuant® WMH</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Manual Tracing</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

Notes: Y = Yes significant findings ($p < .05$); N = No significant findings ($p < .05$); WMH = White matter hyperintensities

Comparison of Total White Matter Volume Methods

Pearson’s correlation comparing NeuroQuant® and FreeSurfer total WM volumes was significant ($r = .38$, $p = .004$, $n = 58$, see Figure 10), with a fair agreement ($ICC [58] = .52$, CI = -.20 to .72). The Bland-Altman plot mean difference was -.007 with 0/58 data points outside the 95% limit of agreement of -.08 to .07 (see Figure 10 for Bland-Altman plot).

Comparison of Total White Matter Volumes by Injury Severity

NeuroQuant® total WM volumes and manual WMH volumes were used for comparisons in H4a and H4b. Two participants did not have NeuroQuant® total WM volumes, resulting in 31 participants with complicated-mild and 27 with moderate-severe TBI. The independent sample t-test found no significant difference in total WM volumes comparing complicated-mild (27.8 ± 2.4) and moderate-severe (28.4 ± 2.5) TBI participants; $t (56) = -.87$, $p = .39$ (see Figure 11). The effect size was quite small (.23).
Comparison of Total Matter Volumes and White Matter Hyperintensities by Injury Severity

There were 19 participants with WMHs and 11 with no WMHs in the complicated-mild TBI group, and 19 participants with WMHs and 7 with no WMHs in the moderate-severe TBI group. The independent sample t-test found no difference in total WMH volumes in participants with WMH (27.64 ± 2.44) and without WMHs (27.99 ± 2.52) in the complicated-mild TBI group; \( t(28) = -.37, p = .71 \). Nor was there a significant difference in total WM volumes comparing participants with (28.22 ± 2.74) and without WMHs (28.85 ± 2.01) in moderate-severe TBI group; \( t(24) = -.55, p = .59 \) (see Figure 11). The effect sizes were small (< .26) for both comparisons.

![Total WM Volumes Between Complicated-Mild and Moderate-Severe TBI, With and Without WMHs](image)
**Figure 11.** Boxplot of total WM volumes for participants with and without WMHs in complicated-mild and moderate-severe TBI groups. There was no significant difference between participants with and without WMHs in either complicated-mild and moderate-severe TBI subgroups; WM = white matter; WMH = white matter hyperintensity; \( * = p < .05 \).

**Research Question: Comparison of Cognitive Function and Methods of Assessing White Matter**

Of the comparisons between WM methods and cognitive functions, Scheltens ratings \( r = -.41, p = .004 \), manual tracing \( r = -.44, p = .002 \), and NeuroQuant® WMH volumes \( r = -.38, p = .01 \) were negatively correlated with processing speed (see Table 8 for a complete table of SEM comparisons). As described in the introduction, the children in the SOBIK investigation also were assessed with several other cognitive measures, the results of which are reported in Table 8. There were no other significant correlations between WM methods and cognitive test scores. When WM outliers were removed from the analysis, there were no significant correlations between WM methods and cognitive test scores (see Table 8).

Sex differences were found between manually traced WMH volumes with processing speed \( p = .000, NFI = .06 \) and NeuroQuant® WMH volumes with processing speed \( p = .000, NFI = .05 \), where males had significant correlations and females did not. The sex differences were not significant when WM outliers were removed from the analysis. See Table 9 for a complete table of chi-squared results for all comparisons.
Table 8

*Strength and Significance of Correlations between White Matter Methods and Cognitive Measures in SEM Model*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Normal Analysis</th>
<th>WM Outliers Excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p</td>
</tr>
<tr>
<td>Scheltens &lt;-&gt; IQ</td>
<td>-.27</td>
<td>.05</td>
</tr>
<tr>
<td>Scheltens &lt;-&gt; PS</td>
<td>-.41</td>
<td>.004*</td>
</tr>
<tr>
<td>Scheltens &lt;-&gt; Language</td>
<td>.20</td>
<td>.14</td>
</tr>
<tr>
<td>Scheltens &lt;-&gt; Attention</td>
<td>-.15</td>
<td>.26</td>
</tr>
<tr>
<td>Manual &lt;-&gt; IQ</td>
<td>-.20</td>
<td>.14</td>
</tr>
<tr>
<td>Manual &lt;-&gt; PS</td>
<td>-.44</td>
<td>.002*</td>
</tr>
<tr>
<td>Manual &lt;-&gt; Language</td>
<td>.13</td>
<td>.35</td>
</tr>
<tr>
<td>Manual &lt;-&gt; Attention</td>
<td>-.06</td>
<td>.65</td>
</tr>
<tr>
<td>NeuroQuant® WMH &lt;-&gt; IQ</td>
<td>-.19</td>
<td>.16</td>
</tr>
<tr>
<td>NeuroQuant® WMH &lt;-&gt; PS</td>
<td>-.38</td>
<td>.007*</td>
</tr>
<tr>
<td>NeuroQuant® WMH &lt;-&gt; Language</td>
<td>.13</td>
<td>.32</td>
</tr>
<tr>
<td>NeuroQuant® WMH &lt;-&gt; Attention</td>
<td>-.07</td>
<td>.60</td>
</tr>
<tr>
<td>FreeSurfer Total WM &lt;-&gt; IQ</td>
<td>-.03</td>
<td>.81</td>
</tr>
<tr>
<td>FreeSurfer Total WM &lt;-&gt; PS</td>
<td>-.06</td>
<td>.67</td>
</tr>
<tr>
<td>FreeSurfer Total WM &lt;-&gt; Language</td>
<td>.18</td>
<td>.18</td>
</tr>
<tr>
<td>FreeSurfer Total WM &lt;-&gt; Attention</td>
<td>-.01</td>
<td>.95</td>
</tr>
<tr>
<td>NeuroQuant® Total WM &lt;-&gt; IQ</td>
<td>-.01</td>
<td>.95</td>
</tr>
<tr>
<td>NeuroQuant® Total WM &lt;-&gt; PS</td>
<td>-.09</td>
<td>.50</td>
</tr>
<tr>
<td>NeuroQuant® Total WM &lt;-&gt; Language</td>
<td>-.17</td>
<td>.23</td>
</tr>
<tr>
<td>NeuroQuant® Total WM &lt;-&gt; Attention</td>
<td>.21</td>
<td>.12</td>
</tr>
</tbody>
</table>

Note. IQ: Intelligence; PS: processing speed; WMH: white matter hyperintensity; WM: white matter; * = significant following the Benjamini-Hochberg test with a false discovery rate set at .05.
Table 9

*Chi-square Difference Test for Sex Differences for the Model and Individual Comparisons between WM Methods and Cognitive Functions, Regular Analysis and Outliers Excluded*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Regular Comparison</th>
<th>WM Outliers Excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>df</td>
<td>p</td>
</tr>
<tr>
<td>Model</td>
<td>20</td>
<td>.03</td>
</tr>
<tr>
<td>Scheltens &lt;&gt; IQ</td>
<td>1</td>
<td>.85</td>
</tr>
<tr>
<td>Scheltens &lt;&gt; PS</td>
<td>1</td>
<td>.01</td>
</tr>
<tr>
<td>Scheltens &lt;&gt; Language</td>
<td>1</td>
<td>.04</td>
</tr>
<tr>
<td>Scheltens &lt;&gt; Attention</td>
<td>1</td>
<td>.50</td>
</tr>
<tr>
<td>Manual &lt;&gt; IQ</td>
<td>1</td>
<td>.23</td>
</tr>
<tr>
<td>Manual &lt;&gt; PS</td>
<td>1</td>
<td>.000*</td>
</tr>
<tr>
<td>Manual &lt;&gt; Language</td>
<td>1</td>
<td>.18</td>
</tr>
<tr>
<td>Manual &lt;&gt; Attention</td>
<td>1</td>
<td>.75</td>
</tr>
<tr>
<td>NeuroQuant® WMH &lt;&gt; IQ</td>
<td>1</td>
<td>.37</td>
</tr>
<tr>
<td>NeuroQuant® WMH &lt;&gt; PS</td>
<td>1</td>
<td>.000*</td>
</tr>
<tr>
<td>NeuroQuant® WMH &lt;&gt; Language</td>
<td>1</td>
<td>.04</td>
</tr>
<tr>
<td>NeuroQuant® WMH &lt;&gt; Attention</td>
<td>1</td>
<td>.59</td>
</tr>
<tr>
<td>FreeSurfer Total WM &lt;&gt; IQ</td>
<td>1</td>
<td>.46</td>
</tr>
<tr>
<td>FreeSurfer Total WM &lt;&gt; PS</td>
<td>1</td>
<td>.81</td>
</tr>
<tr>
<td>FreeSurfer Total WM &lt;&gt; Language</td>
<td>1</td>
<td>.13</td>
</tr>
<tr>
<td>FreeSurfer Total WM &lt;&gt; Attention</td>
<td>1</td>
<td>.34</td>
</tr>
<tr>
<td>NeuroQuant® Total WM &lt;&gt; IQ</td>
<td>1</td>
<td>.82</td>
</tr>
<tr>
<td>NeuroQuant® Total WM &lt;&gt; PS</td>
<td>1</td>
<td>.13</td>
</tr>
<tr>
<td>NeuroQuant® Total WM &lt;&gt; Language</td>
<td>1</td>
<td>.93</td>
</tr>
<tr>
<td>NeuroQuant® Total WM &lt;&gt; Attention</td>
<td>1</td>
<td>.89</td>
</tr>
</tbody>
</table>

Note. IQ: Intelligence; PS: processing speed; WMH: white matter hyperintensity; WM: white matter; * = significant following the Benjamini-Hochberg test with a false discovery rate set at .05.
Discussion

There is considerable interest as well as scientific importance in understanding how well automated MRI methods relate to gold standard research and clinical methods (Azab et al., 2015; Bigler et al., 2010; Caligiuri et al., 2015; Hernández et al., 2017; Kinnunen et al., 2011; Schoemaker et al., 2016; Wenger et al., 2014), as well as how qualitative and quantitative methods of assessing WM relate to cognitive function (Bigler, 2016; Filley & Fields, 2016; Kinnunen et al., 2011; Pantoni, Poggesi, & Inzitari, 2007). This study found significant correlations and good agreement between WMH volumes and WMH ratings. However, no association was found between WMH volumes measured by NeuroQuant® and manual traced volumes. Total WM volumes did not significantly differ by injury severity (mild-complicated vs. moderate-severe TBI) or between participants with and without WMHs present. WMH volumes (NeuroQuant® and manual tracing) and Scheltens WMH ratings were correlated with processing speed. In addition, a sex difference was noted in the correlations between processing speed with manual tracings and NeuroQuant® WMH volumes, with males having faster processing speed and larger WMH volumes compared to females.

Comparison of Quantitative White Matter Hyperintensity Methods

In comparing volumetric methods of assessing WMH, manually traced WMH volumes were not correlated and had poor agreement with NeuroQuant® WMHs volumes (see Table 7). As a reminder, a log transformation was used to adjust for non-normal distributions, which excluded participants who had no identified WMHs (a volume of 0). Thus, the log transformed data used in the ICC and Bland-Altman plot only included those with WMHs identified by both methods, thereby reducing between method variability. As such, good agreement was expected; however, regardless of the log transformation the ICC and Bland-Altman plot indicated poor
agreement between NeuroQuant® and manual tracing WMH volumes. The Bland-Altman plot showed a bias where NeuroQuant® WMH volumes were consistently larger than manual tracing WMH volumes (see Figure 10), suggesting that NeuroQuant® overestimated WMH volumes compared to manual tracing. The finding of consistently larger volumes when using NeuroQuant® may not constitute a major drawback for many studies, if the larger volumes were consistent across individuals and injury severities (Wenger et al., 2014). However, in addition to NeuroQuant consistently overestimating WMH volumes the Bland-Altman plot showed better agreement between larger WMH volumes than for smaller WMH volumes (see Figure 10). Thus volumetric agreement was poor between NeuroQuant® and manual tracings of WMH volumes.

There was no correlation was found between NeuroQuant® and manual tracings WMH volumes. It appears that the lack of correlation is likely due to between method (NeuroQuant® and manual tracing) differences regarding whether WMHs were present or not. Manual tracing identified 18 participants who had no WMHs, whereas NeuroQuant® identified WMHs in all participants. Of the 18 participants that manual tracing identified as having no WMHs, the mean NeuroQuant® WMH volume was .04 cm³, with one outlier at .39 cm³. On closer examination, the outlier appeared to have been a methodological issue of handling artifact, which will be discussed in detail below.

These findings suggest that (1) NeuroQuant® and manually traced WMH methods obtain different results when small or no WMHs are present, and (2) correlation and agreement between the NeuroQuant® and manual tracings improved when larger WMH volumes are present. Similar findings come from a study by Hernández et al. (2017) who also found increased variability between WMH methods for smaller hyperintensities and a stronger, more consistent agreement between methods when larger WMHs were present. Alternatively, strong correlations
(\(r = .98\)) between NeuroQuant® WMH volumes and manually traced WMH volumes were found in patients with multiple sclerosis (Luo et al., 2017b). However, a major difference between findings from the current study and those of Luo et al. (2017b) is that Luo and colleagues only included participants with large WMH volumes (> 5 cm\(^3\)), while the largest WMH volume in the current study was 1.23 cm\(^3\), and the majority of WMHs were small, with some participants having no WMHs (depending on which method was used). Taken together these findings suggest that NeuroQuant® and manual traced WMH volumes could be correlated when WMH volumes are large, but not when WMH volumes are small or absent.

Another possible explanation for the methodological differences between manually traced and NeuroQuant® WMH volumes may be how artifact was handled by each method. In particular, FLAIR sequences used to assess WMHs are sensitive or prone to motion and scanner artifact, especially in parenchymal areas near boney sinuses and fourth ventricle, and choroid plexus (Lisanti et al., 2007; McCarthy et al., 2015; Stuckey, Goh, Heffernan, & Rowan, 2007); see Lisanti et al. 2007 for excellent review of MRI artifact). Since automated methods like NeuroQuant® use thresholding algorithms to identify hyperintensities on FLAIR, they may over categorize scan artifact and incorrectly classify them as WMHs (Wang et al., 2012; Wu et al., 2006). Alternatively, manual tracing may under classify WMH volumes due to raters overgeneralizing the extent of artifact, subsequently missing WMHs. That is raters may incorrectly classify some WMHs as artifact. Short of a post-mortem confirmation, there is not a way to obtain the true anatomical WMH volumes; as such, manual tracings remain the gold standard for volumetric research (Ambarki et al., 2011; Grimm et al., 2015; Lyden et al., 2016; Wenger et al., 2014; Wu et al., 2006). However, the consistency and reproducibility of applied algorithms used by automated methods are often preferred over manual methods as any error is
systematic across subjects rather than individually influenced by raters. In manual tracing the errors are more variable as rater experience, expertise, and bias vary. In this regard, NeuroQuant® may have the advantage of an algorithm that eliminates the variability in rater assessment. While this study found poor agreement and correlation between manual tracings and NeuroQuant® WMH volumes, a recent update of NeuroQuant® (2.3) markets changes to the algorithm that improve lesion detection through a customizable thresholding function, as well as reduce the impact of FLAIR WMH artifact (Harrison, 2017). We used NeuroQuant® 2.0, which does not have these recent improvements. Additional research using the latest NeuroQuant® algorithm is need to assess if manual tracing and NeuroQuant® WMH volumes will be associated using the new version of the software.

**Comparison of Qualitative and Quantitative White Matter Hyperintensity Methods**

Both NeuroQuant® and manually traced WMH volumes correlated with and had good agreement with Scheltens WMH ratings. This finding is similar to previous investigations that found strong correlations between WMH ratings and WMH volumes in healthy aging studies (Gouw et al., 2008; van den Heuvel et al., 2006b; Kapeller et al., 2003). This study is novel as it found a correlation between WMH volumes and ratings in pediatric TBI. Subtle but important methodological differences in classifying WMHs may help explain the good agreement and correlation between Scheltens ratings with manual tracing and NeuroQuant® WMH volumes and the lack of correlation and agreement between manual tracing and NeuroQuant® WMH volumes (see Figure 12). First, out of the three methods manual tracing had more stringent requirements for classifying WMHs including: verifying FLAIR hyperintensities with T2 scans and excluding normal appearing periventricular hyperintensities, which are frequently found on FLAIR images as a result of being located adjacent to CSF filled lateral ventricles, but are not of clinical
importance. (Alam & Sahu, 2010; Lisanti et al., 2007; Zimmerman et al., 1986). As such, the manually traced WMH volume method may be more conservative, or have more requirements to classify WMHs than Scheltens ratings or NeuroQuant® (see figure 12 below), which resulted in manual tracings excluding some WMHs that Scheltens ratings or NeuroQuant® identified.

Additionally, Scheltens rating and NeuroQuant® methods were not designed to determine abnormal or normal periventricular WMHs like manual tracing. Scheltens ratings simply assessed whether periventricular WMHs were present or not and rated the size of the periventricular WMH (ratings between 0-2, with two indicating larger size). NeuroQuant®’s automated algorithm was entirely inclusive of any and all hyperintensities located in WM, including periventricular regions and artifact. As such, Scheltens and NeuroQuant® may be more liberal, or more inclusive in its classification of WMHs than manual tracing, with NeuroQuant® being the most liberal, or inclusive of the three WMH methods. Finally, because Scheltens rating and manual tracing are both rater based methods capable of visually separating artifact from WMHs, Scheltens is slightly more conservative than NeuroQuant® but not to the point of disagreement. Scheltens, manual tracing, and NeuroQuant® are comparable methods.

Further evidence was seen in the number of individuals identified with no WMHs by each method: manual tracing identified 18 participants, Scheltens ratings identified 7 participants, and NeuroQuant® identified all participants as having some WMH (no participant identified with no WMHs). Similar findings were found by Ross and colleagues (2015) using an older version of NeuroQuant® which identified more pathology than board certified radiologists in adult TBI. In the present study, NeuroQuant® identified no individuals without WMHs; however, Scheltens ratings identified 7 participants who did not have WMHs. The agreement between Scheltens ratings with manual tracing and NeuroQuant® WMH volumes may be due to...
the fact that Scheltens ratings identified WMHs in most individuals and fell between manual tracing and NeuroQuant® (see Figure 12). Whereas manual tracing and NeuroQuant® differed in which participants had WMHs or no WMHs, as such these methods were not correlated or agreed. Caution should be used when comparing manual tracings with NeuroQuant® 2.0 WMH volume findings.

<table>
<thead>
<tr>
<th><strong>More Conservative</strong></th>
<th><strong>Moderate</strong></th>
<th><strong>More Liberal</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Excludes normal appearing periventricular WMHs.</td>
<td>• Rates all periventricular regions as present or absent.</td>
<td>• Includes any and all hyperintensity located in WM, which may include some artifact.</td>
</tr>
<tr>
<td>• Raters have the ability to determine if WM abnormalities are artifact or WMHs</td>
<td>• Raters have the ability to determine if WM abnormalities are artifact or WMHs</td>
<td></td>
</tr>
<tr>
<td>• Used T2 image to confirm WMH on FLAIR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Number of SOBIK Participants with no identified WMHs</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• 18 participants</td>
</tr>
</tbody>
</table>

*Figure 12.* Descriptive summary of methodological differences for identifying WMHs between methods. Manual tracings had the most conservative requirements for identifying WMHs, whereas NeuroQuant® had the most liberal requirements for identifying WMHs. Evidence for the classification of methodological differences between WMH methods can be seen via the number of participants identified with no WMHs.
Interestingly, all the WMHs were correlated with processing speed regardless of method used (see discussion below), suggesting that these methods can be used in both clinical and research settings. In the end, the decision regarding which method to use may depend on time, funding, access to trained raters, availability of technical programs and computational power, and purpose of assessing WMHs. The increasing use of automated neuroimaging segmentation methods (e.g., FreeSurfer and NeuroQuant®) in published articles is apparent through a simple scholarly review (see Figure 13). Automated methods are often preferred because of their fast, consist, quantitative results. Manual tracing requires extensive training, technical programs, time, expertise and training, and are subject to rater bias, thereby decreasing accessibility and utility in clinical settings. Many of automated methods have been validated with manual tracings, including NeuroQuant® as previously discussed (Kovacevic et al., 2009; Lyden et al., 2016), and FreeSurfer (Fischl, 2012; Fischl et al., 2002). Scheltens ratings are a simple, inexpensive method and can be completed with minimal training and quickly performed with no other requirement than being able to visually inspect the image. NeuroQuant® costs ~$89 per report, whereas there are no costs for using the Scheltens rating scale, except for the time to visually inspect the image. An important caveat is that Scheltens rating scale does not provide volumetric information, which limits the ability to track changes with precision over time (i.e., following therapy, recovery, or atrophy; Caligiuri et al., 2015; Gouw et al., 2008; van den Heuvel et al., 2006b; Kapeller et al., 2003). Nonetheless, Scheltens rating scales can be used in MRI studies with motion, bone or other artifacts, where automated methods are somewhat limited, or may not run at all when there is artifact.
Figure 13. A search of PubMed using search terms: “NeuroQuant”, “LesionQuant”, “Scheltens rating”, “manual tracing”, and “FreeSurfer” was conducted. Published articles between 1984 and 2017 show a significant increase in published studies that use automated imaging methods. Manual methods still remain the gold standard in research. The labels indicate the creation date of each method.

Comparison of Total WM Volume Methods

In comparing total WM methods, NeuroQuant® and FreeSurfer had significant correlation and fair agreement. Because NeuroQuant®’s segmentation algorithm originated from FreeSurfer (Kovacevic et al., 2009; Ochs et al., 2015), the correlation and fair agreement were expected to be stronger. Similar to Ochs et al. (2015), these results showed that NeuroQuant®’s total WM volumes were slightly larger than those obtained by FreeSurfer. Importantly, this is the
first study to examine this comparison in a pediatric sample. Previous studies comparing an earlier version of NeuroQuant® and FreeSurfer found good association in healthy adult individuals (Zhou, Zambelli, & Heredia, 2015), multiple sclerosis (Nielsen & Kita, 2017), and following TBI (Reid et al., 2017). The large confidence interval for the ICC reduce the confidence in this finding. The large confidence interval (-.20 to .72) between NeuroQuant® and FreeSurfer includes the zero or null value which means there is no statistically meaningful difference between the two methods (Cicchetti & Sparrow, 1981). One possible explanation for the wide confidence interval may be the different atlases used between this version of NeuroQuant (2.0, dynamic atlas) and FreeSurfer, which until this study have not been compared. Additional research is needed to compare total WM volumes using NeuroQuant® and FreeSurfer in pediatric TBI.

**Comparison of Total White Matter Volumes by Injury Severity**

Differences in total WM volumes based on injury severity have been reported following TBI (Beauchamp et al., 2011; Bigler et al., 2013a), especially between mild and severe TBI. In this study, there was no significant difference in total WM volumes when comparing participants with complicated-mild to moderate-severe TBI (see Figure 11), and the effect size was small. This may be due to the fact that the children all attended regular school suggesting good recovery. However, this study was under powered to detect such differences and there was not a comparison to healthy controls. A post-hoc power analysis using the ClinCalc calculator (http://clincalc.com/stats/power.aspx) was carried out using an alpha of .05, resulting in 13.9% power. To achieve sufficient statistical power (80% power), an additional 270 participants would be needed to detect between group differences.
Even mild-complicated TBI is likely resulting in morphometry changes, but a most recent study by Dennis et al. (2017) has provided additional insight why there was not an injury severity differences in WM volume. In the Dennis et al. (2017) study they examined volumetric findings in children with moderate-to-severe TBI who had negative versus positive outcomes. It was only the participants with negative outcomes who had WM volumetric loss over time (children were assessed at 2-5 months and 12 months post-injury). Wu et al. (2010) also demonstrated that over time there was a reconstitution of white matter integrity in some children with TBI. The children in the SOBIK investigation were a select sample as they had to have achieved a level of recovery that allowed them to be in a regular classroom with minimal to no accommodation. Accordingly, this pediatric TBI sample is not representative of the full spectrum of TBI. This likely partially explains why there was not a significant difference related to the severity of injury. In addition, the small WMH volumes identified in this sample may have been too small to significantly affect total WM volumes. It could also be that the FLAIR identified WMH volumes were included as part of the total WM volumes on T1 scans as hypointensities on T1 often underrepresent FLAIR identified WMHs (see Figure 2) (Bigler, 2015; Sanfilipo et al., 2005).

Using the SOBIK data, Bigler et al. (2013a) found a significant difference in total WM volumes between complicated-mild and moderate-severe TBI using FreeSurfer total WM volumes. In the current study, NeuroQuant® total WM volumes were used. In order to compare our results to those previously published by Bigler and colleagues, the analysis was redone using the FreeSurfer data and confirmed the Bigler et al. findings. As this study showed FreeSurfer and NeuroQuant® total WM volumes were correlated with good agreement, this differing finding was unexpected. A post hoc analysis was performed in the current study comparing
NeuroQuant® and FreeSurfer total WM volumes by injury severity (complicated-mild, moderate, and severe). Total WM volumes differed significantly between NeuroQuant® and FreeSurfer in complicated-mild TBI group ($p = .02$) but not the moderate or severe TBI group. The lack of difference in the moderate to severe group may be due to small sample size in the moderate to severe group ($n = 19$), or it may represent subtle differences found between the dynamic atlas of NeuroQuant® and static atlas of FreeSurfer. Ultimately the reason is unclear and requires additional study.

**Comparison of Total White Matter Volumes and White Matter Hyperintensities**

This is the first study to compare WMHs with total WM volumes. There were no differences in total WM volumes when comparing participants with and without WMHs in complicated-mild or moderate-severe TBI groups. Even in the presence of a WMH, the tissue classification algorithms that separate WM volume from GM do not isolate WMH as being different or distinct from overall total WM volume. When the analysis was re-run combining both the complicated-mild and moderate-severe TBI groups, the results did not change. While WM atrophy is common, studies in pediatric TBI have shown areas of increased volumes, including in WM regions (Bigler, 2016; Dennis et al., 2016). In the chronic phase of pediatric TBI, Dennis et al. (2016) found WM atrophy, but also found increased WM volumes in five WM regions including the angular gyrus, occipital gyrus, posterior thalamic radiation, superior longitudinal fasciculus, and superior parietal lobule. This increase may be due to compensatory mechanisms and rewiring of the brain following TBI. As such, assessing WM integrity through total WM volumes may be insufficient to capture the distinct pathological changes following pediatric TBI. Additionally, the lack of relationship between WMH and total WM volumes may be due to the relatively small sample resulting in low power. A post-hoc power analysis was
carried out using the ClinCalc calculator (http://clincalc.com/stats/power.aspx). With an alpha of .05, the comparison with complicated-mild TBI had a power of 5.6%, and required an additional 770 participants to achieve sufficient statistical power (80% power) to detect between group differences. The comparison with moderate-severe TBI had a power of 9.3%, and required an additional 213 participants achieve sufficient statistical power (80% power) to detect between group differences. Larger samples that include a more complete range of TBI severity is needed to add support to this finding.

**Research Question: Comparison of Cognitive Function and Methods of Assessing White Matter**

When cognitive scores were separated by injury severity, the moderate-severe TBI group had higher cognitive scores than did the complicated-mild TBI group for most cognitive scores (see Table 6). When moderate-severe TBI was separated into moderate (n = 9) and severe (n = 19) TBI groups, the moderate TBI group had the highest scores (better scores) than did the complicated-mild and severe subgroups. The severe TBI group had the lowest scores, as expected. When the scores were of the moderate and severe groups were averaged together, the combined moderate-severe TBI group had higher cognitive scores than did the complicated-mild group, as such these findings should be interpreted cautiously.

**Processing speed.** From the SEM model, the Scheltens ratings, manually traced WMH volumes, and NeuroQuant® WMH volumes significantly correlated with processing speed. There were no other significant correlations between WM volumes or ratings and any other cognitive domain. Processing speed may be the most directly related to WM integrity in TBI. Numerous studies have shown that processing speed is associated with WM integrity (Anderson et al., 2012; Bigler et al., 2013a; Gene et al., 2017; Kochunov et al., 2010; Kourtidou et al., 2013;
Shenton et al., 2012; Wozniak et al., 2007) and WMHs (Junqué et al., 1990; Papp et al., 2014; Vannorsdall et al., 2009; Wakefield et al., 2010). The association between WMHs and processing speed occur in a variety of populations including healthy elderly adults (Debette & Markus, 2010; Königs et al., 2018; Kuznetsova et al., 2016; Papp et al., 2014; van den Heuvel et al., 2006a; Vannorsdall et al., 2009; Vernooij et al., 2009), in individuals with Type 2 diabetes (Reijmer, Leemans, Brundel, Kappelle, & Biessels, 2013), cerebral small vessel disease (Duering et al., 2014; Righart et al., 2013), multiple sclerosis (Papadopoulou et al., 2013; Sperling et al., 2001), and children with sickle cell disease (van der Land et al., 2015). Further, a meta-analysis of studies focusing on adults from 1966 to 2009 found slow processing speed was associated with WMHs (Debette & Markus, 2010). Regardless of the etiology of WMHs, there is strong evidence of the association between WMHs and slow processing speed suggesting that WMHs may be a structural marker of impaired processing speed.

While there were strong associations between WMH volumes and processing speed there was no relationship between total WM volumes and processing speed. Bigler et al. (2013a) also found a lack of relationship between total WM volumes and processing speed. Magistro et al. (2015) found that processing speed was associated with total WM volumes; however, Magistro’s participants were healthy young adults without brain injury. As previously discussed, there are also compensatory mechanisms following TBI, such as increased WM volume in certain areas (Bigler, 2016; Dennis et al., 2016), which may disguise WM pathology in assessments using total WM volume, which may contribute the lack of an association with processing speed.

**Intelligence.** There were no associations between total WM and WMH methods and intellectual functioning. The link between WM and intellectual function after pediatric TBI is a point of discussion in the literature. Some studies found a relationship between WM volumes and
intelligence (Anderson et al., 2012; Königs et al., 2018; Wozniak et al., 2007), while others find no relationship (Anderson et al., 2006; Catroppa, Anderson, Morse, Haritou, & Rosenfeld, 2008; Gale & Prigatano, 2010). Anderson et al. (2012) found that total WM volumes did not predict intelligence 10 years post injury. Alternatively, Gale and Prigatano (2010) found that WM volume loss in pediatric TBI children was associated with WISC-IV Full Scale IQ scores an average of 3.3 years post injury.

In this study, all SOBIK participants attended regular classrooms and most had normal or average intellectual function, at least at a group level (Mean = 99, SD = 15.10). This means that children with significant intellectual disabilities, or individuals needing remedial or exclusive special education as a result of their TBI, were not included in this study. Average intelligence and an ability to reintegrate back into a normal classroom setting may suggest that participants had higher cognitive reserves pre injury. Cognitive reserve has been shown to be a buffer for the negative cognitive effects of brain injury (Benedict, Morrow, Guttman, Cookfair, & Schretlen, 2010; Brickman et al., 2011; Leary et al., 2018), and may explain the generally intact intellectual and cognitive function in the SOBIK participants.

It is unclear if the TBI participants who attend regular school will continue at the same level of intellectual function as they age. Anderson et al., (2012), assessed cognitive function ten years after TBI and found that pediatric TBI survivor’s intelligence declined to the average to low average range regardless of injury severity. In addition, children who had a severe TBI prior to age 7 often failed to keep pace with normal intellectual development (Anderson et al., 2012). Participants in this study were an average of 2.6 years post injury. Additional testing at a later time is needed to determine if their intellectual function will decline over time.
Attention. There were no significant associations between WM measures and attention. Similar to these findings, a study by Power, Catroppa, Coleman, Ditchfield, and Anderson (2007) found neither the location nor severity of WM lesions predicted attentional deficits. There is evidence from DTI studies which show lower FA values were associated with attention deficits (Kraus et al., 2007; Wozniak et al., 2007). The reason for the lack of association between WM volumes and attention is unclear, but may be due to small, relatively high functioning sample or because attention was assessed by averaging scores across several subtests.

Language. There were no significant associations between total WM volumes and language. Gale & Prigatano (2010) found that WM volume loss was associated with language difficulties (i.e., verbal fluency), while Königs et al. (2018) found no associations. The location of WM injury may be a better predictor of language problems than the extent of the pathology. For example, damage to the language areas in the frontal and temporal lobes or to the arcuate fasciculus (Liégeois et al., 2013), may result in impairments in language. The current study assessed WMH and total WM volumes rather than lesion location, and given the small size/volume of WMH lesions in the sample, the lack of association with language was not surprising. In addition, the assessment of language was extremely limited in the SOBIK study as the only measure of language was the Recreating Speech Acts subtest of the TLC-E (Wiig & Secord, 1985). Finally, participants were excluded from the SOBIK study if they had obvious language impairments.

Sex Differences

Males and females differed in their comparisons between manual tracing WMH volumes and processing speed and NeuroQuant® WMH volumes with processing speed, where males had significant correlations and females did not. This was not unexpected as males had larger
WMHs than females and a sex differences in processing speed are well documented with males performing slightly better than females (Strauss et al., 2006). Interestingly, the TEA-Ch has been shown to have sex differences for the Creature Counting subtest (Manly et al., 1999; Strauss et al., 2006), which was not identified in this study. It is possible that the difference between males and females, if it existed, was obscured when the three subtests from the TEA-Ch were averaged, but there is no data to support this from the SEM model.

**Study Strengths and Limitations**

There were several strengths of this study. The methodological comparisons in this study fill an important gap in the literature, as methods previously validated for use in adults now have supporting evidence for their use in pediatric TBI. This is especially important as NeuroQuant® has marketed it’s expanded age range to include pediatrics (Luo, Airriess, & Albright 2015; White et al., 2015a), but currently has no published articles assessing it’s use within this age group. This is the first study to compare NeuroQuant® to existing clinical and research methods in pediatric TBI. This is also the first study to show comparability between total WM volumes from NeuroQuant® and FreeSurfer following the addition of the dynamic atlas in NeuroQuant® 2.0, showing that regardless of various updates, these methods remain comparable.

Prior to this study, Scheltens ratings had not been compared to WMH volumes in pediatric TBI. This study shows supporting evidence of Scheltens comparability to volumetric methods of WMHs, reassuring its use in both clinical and research settings. While both total WM volumes and WMHs methods have been used in quantitative MRI literature to assess WM integrity, this is the first study to compare their findings to assess potentially redundant information about WM integrity. While this analysis was limited by a small sample size and power, these preliminary findings support the use of WMH methods over total WM volumes to
assess WM integrity. Additionally, in the assessment of cognitive functioning following pediatric TBI, this sample focused on a group of children who attended regular school with chronic TBI, which is widely underrepresented in the literature. This study’s data support prior evidence of the association between processing speed and WM integrity in a unique assessment that compared multiple assessments of WM and cognitive functions.

Although the SOBIK study is one of the largest pediatric investigations when it originally began back in 2006, this study nonetheless is limited by small sample size, which consequently limited statistical power in some of the analyses. Replication in a larger sample is needed. The 1.5 Tesla scans used allowed comparisons of all WM methods and are widely used in the literature to study pediatric TBI (Beauchamp et al., 2011; Bigler et al., 2010; Roberts et al., 2016; Verger et al., 2001; Wu, 2011); however, replication with 3 Tesla scans is needed as 3 Tesla scans have increased resolution which improve identification of WM damage (Neema et al., 2009). Another limitation is that the parent study used an abbreviated neuropsychological test battery, which restricted the assessment of cognitive function because the focus was on social outcome and classroom behavior. A more comprehensive battery may be more sensitive in identifying subtle cognitive deficits.

While we were unable to show good agreement between NeuroQuant® and manual tracing WMH volumes, as previously mentioned, future studies comparing the most recent update of NeuroQuant®, with its marketed improvement on the thresholding of WMHs and handling of artifact, may show improved correlation and agreement between NeuroQuant® and gold standard manual tracings. This would be an important area of investigation in assessing the reliability. Future examination within larger, more representative samples of TBI are need to confirm these findings.
Conclusions

Given the high rate of TBI during childhood and its effect on cognitive and behavioral functions such as processing speed, methods aimed at identifying and quantifying WM pathology are important clinical and research concerns. Reliable methods of evaluating WM is central to aid in diagnosis, assess recovery and/or changes over time, and provide objective evidence to support clinical observations and impressions as well as treatment (Brewer, 2009; Murphy & Duhaime, 2016). These results provide support for several conclusions regarding the comparability of WMH methods, total WM volume comparisons, and their associations with cognitive function in pediatric TBI. First, NeuroQuant® and manually traced WMH volumes do not correlate or agree with each other, with NeuroQuant® having consistently larger volumes of WMH compared to manual tracing WMH volumes. Second, Scheltens WMH ratings were associated with both NeuroQuant® and manually traced WMH volumes. Third, total WM volumes did not adequately capture WM pathology as identified by WMHs, thus findings between total WM volumes and WMHs differ. Fourth, processing speed was associated with WMH methods (i.e., Scheltens ratings, manual tracings, NeuroQuant® WMH volumes) but not total WM methods (i.e., FreeSurfer and NeuroQuant®), suggesting that WMH methods can be used in clinical and research settings to provide support in assessments of processing speed. No other significant association between WM methods and cognitive functions were observed. Finally, males had significant correlations between manual tracings and NeuroQuant® WMH volumes with processing speed, and females did not. This may be because males have larger WMH volumes than females. In conclusion, WMH methods can be used in clinical and research settings to provide support on assessments of processing speed. The decision as to which method
to use to assess WM will likely depend on available resources, time, and purpose of assessment or study design.
References


doi:10.1177/0883073816654143


during the first year after moderate to severe traumatic brain injury. *Neuroimage, 5*, 128-140. doi:http://dx.doi.org/10.1016/j.nicl.2014.03.012


Duering, M., Gesierich, B., Seiler, S., Pirpamer, L., Gonik, M., Hofer, E., . . . Dichgans, M.
(2014). Strategic white matter tracts for processing speed deficits in age-related small

Youth risk behavior surveillance -- United States, 2009. Retrieved from
https://www.cdc.gov/mmwr/preview/mmwrhtml/ss5905a1.htm

White matter hyperintensities and their associations with suicidality in psychiatrically
hospitalized children and adolescents. *Journal of the American Academy of Child &
Adolescent Psychiatry, 43*(6), 770-776. doi:10.1097/01.chi.0000120020.48166.93

Engedal, K., Braekhus, A., Andreassen, O. A., & Nakstad P. H. (2012). Diagnosis of dementia-
doi: 10.4045/tidsskr.12.0148

resonance images. *Magnetic Resonance in Medicine, 44*(6), 973-977. doi:
https://doi.org/10.1002/1522-2594(200012)44:6<973::AID-MRM21>3.0.CO;2-H

and variation coefficients for log transformed variables. *Journal of Clinical
Epidemiology, 61*(10), 978-982. doi: https://doi.org/10.1016/j.jclinepi.2007.11.003


  doi:10.1016/j.neuroimage.2012.01.021


doi:https://doi.org/10.1017/S1355617700000060


Kochanek, P. M., Clark, R. S. B., Ruppel, R. A., Adelson, P. D., Bell, M. J., Whalen, M. J., . . .


Kochunov, P., Coyle, T., Lancaster, J., Robin, D. A., Hardies, J., Kochunov, V., . . . Fox, P. T.

(2010). Processing speed is correlated with cerebral health markers in the frontal lobes as

doi:10.1016/j.neuroimage.2009.09.052

doi:10.1111/ene.12719


doi:http://dx.doi.org/10.1016/j.jcm.2016.02.012


doi:10.1097/WAD.0b013e318192e745


neuropsychological outcomes following traumatic brain injury. *Journal of Head Trauma Rehabilitation, 33*(1), E28-E35. doi:10.1097/HTR.0000000000000329


doi:http://dx.doi.org/10.1016/S0925-4927(01)00129-9


doi:10.1080/01688639208402865


doi:10.1002/ana.22455


doi:10.1191/026921598672178340


doi:https://doi.org/10.1016/j.brainres.2007.04.002


doi:10.1016/j.expneurol.2005.07.014


doi:https://doi.org/10.1080/02699052.2017.1341645


doi:10.1080/13803395.2015.1123225

resonance imaging. *Journal of the Neurological Sciences, 114*(1), 7-12. doi:10.1016/0022-510X(93)90041-V


reliability and validity issues in younger and older brains. *Human Brain Mapping*, 35(8), 4236-4248. doi:10.1002/hbm.22473


doi:10.1097/01.chi.0000159947.50523.b4


doi:10.1017/S1355617712001531

doi:10.10170S1355617704103093


doi:10.1016/j.neuroimage.2009.02.004


Appendix A: Example of NeuroQuant®’s Multi Structure Atrophy Report

**NeuroQuant®**
Multi Structure Atrophy Report

**PATIENT INFORMATION**

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Patient Name</th>
<th>Sex</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1104</td>
<td>1104</td>
<td>M</td>
<td>12</td>
</tr>
<tr>
<td>Accession Number</td>
<td>Referring Physician</td>
<td>Exam Date</td>
<td></td>
</tr>
<tr>
<td>1104</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**MORPHOMETRY RESULTS (1 of 2)**

<table>
<thead>
<tr>
<th>Brain Structure</th>
<th>Volume (cm³)</th>
<th>% of ICV (5%-95% Normative Percentile)</th>
<th>Normative Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole Brain</td>
<td>1280.93</td>
<td>82.94 (76.50 - 84.59)</td>
<td>82</td>
</tr>
<tr>
<td>Lateral Ventricles</td>
<td>17.59</td>
<td>1.14 (0.48 - 1.72)</td>
<td>74</td>
</tr>
<tr>
<td>Thalami</td>
<td>18.13</td>
<td>1.17 (0.99 - 1.20)</td>
<td>91</td>
</tr>
</tbody>
</table>

**AGE-MATCHED REFERENCE CHARTS**

- L & R Whole Brain
- L & R Lateral Ventricle
- L & R Thalamus

CorTechs Labs, Inc. | cortechslabs.com
COMPARISON OF WM METHODS IN PEDIATRIC TBI

NeuroQuant®
Multi Structure Atrophy Report

PATIENT INFORMATION
Patient ID: 1104
Patient Name: 1104
Sex: M
Age: 12

MORPHOMETRY RESULTS (2 of 2)

<table>
<thead>
<tr>
<th>Brain Structure</th>
<th>Volume (cm³)</th>
<th>% of ICV</th>
<th>5%-95% Normative Percentile</th>
<th>Normative Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortical Gray Matter</td>
<td>559.19</td>
<td>36.21</td>
<td>34.85 - 43.29</td>
<td>12</td>
</tr>
<tr>
<td>Cerebral White Matter</td>
<td>490.73</td>
<td>31.77</td>
<td>23.88 - 30.07</td>
<td>&gt; 99</td>
</tr>
<tr>
<td>White Matter Hypointensities*</td>
<td>0.22</td>
<td>0.01</td>
<td>0.00 - 0.07</td>
<td>52</td>
</tr>
<tr>
<td>Third Ventricle</td>
<td>0.85</td>
<td>0.05</td>
<td>0.02 - 0.08</td>
<td>59</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>8.71</td>
<td>0.56</td>
<td>0.47 - 0.63</td>
<td>52</td>
</tr>
<tr>
<td>Inferior Lateral Ventricles</td>
<td>1.08</td>
<td>0.07</td>
<td>0.04 - 0.11</td>
<td>55</td>
</tr>
</tbody>
</table>

AGE-MATCHED REFERENCE CHARTS

*White matter hypointensities are abnormally low signal intensity regions within white matter as observed on a T1-weighted MRI scan.
Figure 14. Example of the Multi Structural Atrophy Report from NeuroQuant®, which used T1 images to quantify whole brain, lateral ventricles, thalamus, cortical gray matter, cerebral white matter, white matter hypointensities, third ventricle, hippocampi, and inferior lateral ventricle volumes. Volumetric data is compared to normal age-matched control volumes for each region of interest, represented as a normative percentile.
Appendix B: Example of NeuroQuant®’s LesionQuant Report

PATIENT INFORMATION

<table>
<thead>
<tr>
<th>Patient ID:</th>
<th>Patient Name:</th>
<th>Sex:</th>
<th>Age:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1104</td>
<td>1104</td>
<td>M</td>
<td>12</td>
</tr>
<tr>
<td>Accession Number:</td>
<td>Referring Physician:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1104</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SCANNER INFORMATION

<table>
<thead>
<tr>
<th>Current Scan Date:</th>
<th>Current Scanner Name:</th>
<th>Prior Scan Date:</th>
<th>Prior Scanner Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SIGNA EXCITE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MORPHOMETRY RESULTS

<table>
<thead>
<tr>
<th>Brain Structure</th>
<th>Volume (cm³)</th>
<th>% ICV</th>
<th>Normative Percentile</th>
<th>Volume (cm³)</th>
<th>% ICV</th>
<th>Normative Percentile</th>
<th>Volume Change (cm³)</th>
<th>Volume Change (%)</th>
<th>Norm Percentile Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole Brain</td>
<td>1280.71</td>
<td>82.93</td>
<td>82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral Ventricles</td>
<td>17.78</td>
<td>1.15</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalami</td>
<td>18.14</td>
<td>1.17</td>
<td>91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortical Gray Matter</td>
<td>556.39</td>
<td>36.03</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebral White Matter</td>
<td>487.56</td>
<td>31.57</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third Ventricle</td>
<td>0.83</td>
<td>0.05</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hippocampi</td>
<td>8.72</td>
<td>0.56</td>
<td>62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior Lateral Ventricles</td>
<td>1.07</td>
<td>0.07</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

LESION RESULTS (from current scan)

<table>
<thead>
<tr>
<th>All Lesions from Current Scan</th>
<th>All Lesions from Prior Scan</th>
<th>Enlarging Lesions</th>
<th>New Lesions</th>
<th>New + Enlarging Lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Volume (cm³)</td>
<td>6.06</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>% ICV</td>
<td>0.39</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Lesion Burden</td>
<td>1.24</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

LESION ANATOMICAL DISTRIBUTION (from current scan)

<table>
<thead>
<tr>
<th>Leukocortical</th>
<th>Periventricular</th>
<th>Infratentorial</th>
<th>Deep White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesion Count</td>
<td>10</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>New Lesion Count</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Enlarging Lesion Count</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>New + Enlarging Lesion Count</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lesion Volume (cm³)</td>
<td>4.47</td>
<td>1.26</td>
<td>0.00</td>
</tr>
</tbody>
</table>
COMPARISON OF WM METHODS IN PEDIATRIC TBI

PATIENT INFORMATION

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Patient Name</th>
<th>Sex</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1104</td>
<td>1104</td>
<td>M</td>
<td>12</td>
</tr>
</tbody>
</table>

AGE AND GENDER MATCHED REFERENCE CHARTS (from current scan)
Figure 15. Example of NeuroQuant®’s LesionQuant report, which used FLAIR images to quantify white matter hyperintensities (WMHs) in various brain regions. Anatomical distribution of the hyperintensities are assessed, as well as the number of lesions, the volume, the corrected volume for total intracranial volume, and lesion burden are calculated. Volumetric data is compared to normal age-matched control volumes for each region of interest, represented as a normative percentile.
Appendix C: Description of Neuropsychological Instruments

Processing Speed

Processing speed was assessed via the Processing Speed Index (PSI) from the Wechsler Intelligence Scale for Children- Fourth Edition (WISC-IV; Wechsler, 2003b). The PSI measures the speed of mental and graph-o-motor processing speed (Strauss et al., 2006) and is a valid measure of processing speed in children with TBI (Genc et al., 2017; Shultz et al., 2016; Yeates & Donders, 2005). The PSI includes Coding and Symbol Search tasks, along with the Cancellation subtest. Although PSI scores typically consist of the sum of the scaled scores for the Coding and Symbol Search tasks, the Cancellation subtest can be used as substitution for either the Coding or Symbol Search (Strauss et al., 2006) which was done in the SOBIK study. The Symbol Search task requires scanning a group of symbols and indicate the presence or absence of a target symbol. The Cancellation task requires the participant to scan both a random and nonrandom assortment of pictures and mark a target picture within a given time limit. Time and accuracy were recorded for both measures which was consolidated into a composite age corrected PSI score (Strauss et al., 2006; Wechsler, 2003a). Standardized scores on the PSI range from 40 to 160 with higher scores indicating faster processing speed, with a mean score of 100 and SD of 15.

Intelligence

The Vocabulary and Matrix Reasoning subtests of the Wechsler Abbreviated Scale of Intelligence (WASI) were administered, from which a full scaled intelligence quotient score was attained (Strauss et al., 2006; Wechsler, 1999). The WASI is designed as a brief measure of intelligence using verbal and non-verbal abilities (Genc et al., 2017). The WASI scores are converted to standardized scores, which have a mean of 100 and a standard deviation (SD) of 15,
where higher scores indicated better cognitive function. The WASI has been validated for use in pediatric TBI populations (Bigler et al., 2013a; Genc et al., 2017; Strauss et al., 2006; Wechsler, 1999).

**Attention**

The Test of Everyday Attention in Children (TEA-Ch; Manly et al., 1999) is a widely used and valid measure of attention in children with TBI (Anderson et al., 1998; Araujo et al., 2017; Manly et al., 1999; Strauss et al., 2006). Three subtests were administered. The control/switching attention was measured by the Creature Counting subtest which requires a participant to count creatures, switching from counting forward and backward as they encounter the appropriate signaling arrows (down or up arrows). Sustained/response inhibition was measured by the Walk Don’t Walk task, which requires the participant to mark footsteps along a path with a “go” tone, and stop marking the footsteps with a “stop” tone. Sustained attention is measured by the Code Transmission task, where the participant is instructed to list the numbers that proceeded a target (e.g., two 5s in a row), which occurs infrequently. There is no composite score for the TEA-Ch. Instead, scaled scores are reported for each subtest and have a mean of 10 with a standard deviation of three where higher scores indicated better attention abilities.

**Language**

The Test of Language Competence-Expanded Edition assesses delays in linguistic competence and detects language delays in children (Wiig & Secord, 1985). The SOBIK dataset administered the Recreating Speech Acts subtest, which assesses the ability to produce contextually appropriate verbal responses using key words related to a situation or context (Yeates et al., 2004), to verify that the recruited TBI participants did not have language difficulties. The reported scaled scores have a mean of 10 and a standard deviation of three
where higher scores indicated better language abilities.
Appendix D: Full Scheltens Rating Scale

**Periventricular (0-6)**
- Frontal caps (0-2)
- Occipital caps (0-2)
- Bands (0-2)

**White Matter (0-24)**
- Frontal (0-6)
- Parietal (0-6)
- Occipital (0-6)
- Temporal (0-6)

**Basal Ganglia (0-30)**
- Caudate nucleus (0-6)
- Putamen (0-6)
- Globus pallidus (0-6)
- Thalamus (0-6)
- Internal capsule (0-6)

**Infratentorial (0-24)**
- Cerebellum (0-6)
- Mesencephalon (0-6)
- Pons (0-6)
- Medulla (0-6)

**Key (0-6)**
- 0 = absent
- 1 = ≤ 5 mm
- 2 = > 5 mm and < 10

**Key (0-78)**
- 0 = absent
- 1 = < 3 mm, n ≤ 5
- 2 = <3 mm, n > 6
- 3 = 4-10 mm, n ≤ 5
- 4 = 4 mm, n > 6
- 5 = > 11 mm, n > 1
- 6 = confluent

**Total Score 0-84**

*Figure 16.* Visual of the full Scheltens rating scale used to assess hyperintensities on T2 images;

The modified scale for assessing WMHs excludes all GM regions. The blue box shows the periventricular ratings, which are assessed using a scale of 0-2, where the orange boxes show the WM, basal ganglia, and infratentorial regions, which are assessed using a scale of 0-6.
### Appendix E: Example of Scheltens Ratings for White Matter Regions

#### Periventricular Hyperintensities

<table>
<thead>
<tr>
<th>Region</th>
<th>Scale Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caps occipital</td>
<td>0 = absent&lt;br&gt;1 = ≤ 5 mm&lt;br&gt;2 = &gt; 5 mm and &lt; 10 mm</td>
</tr>
<tr>
<td>Caps frontal</td>
<td>0 = absent&lt;br&gt;1 = ≤ 5 mm&lt;br&gt;2 = &gt; 5 mm and &lt; 10 mm</td>
</tr>
<tr>
<td>Bands lateral ventricles</td>
<td>0 = absent&lt;br&gt;1 = ≤ 5 mm&lt;br&gt;2 = &gt; 5 mm and &lt; 10 mm</td>
</tr>
</tbody>
</table>

#### White Matter Hyperintensities

<table>
<thead>
<tr>
<th>Region</th>
<th>Scale Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frontal</td>
<td>0 = absent&lt;br&gt;1 = &lt; 3 mm, n ≤ 5&lt;br&gt;2 = &lt;3 mm, n &gt; 6&lt;br&gt;3 = 4-10 mm, n ≤ 5&lt;br&gt;4 = 4 mm, n &gt; 6&lt;br&gt;5 = &gt; 11 mm, n &gt;1&lt;br&gt;6 = confluent</td>
</tr>
<tr>
<td>Parietal</td>
<td>0 = absent&lt;br&gt;1 = &lt; 3 mm, n ≤ 5&lt;br&gt;2 = &lt;3 mm, n &gt; 6&lt;br&gt;3 = 4-10 mm, n ≤ 5&lt;br&gt;4 = 4 mm, n &gt; 6&lt;br&gt;5 = &gt; 11 mm, n &gt;1&lt;br&gt;6 = confluent</td>
</tr>
<tr>
<td>Occipital</td>
<td>0 = absent&lt;br&gt;1 = &lt; 3 mm, n ≤ 5&lt;br&gt;2 = &lt;3 mm, n &gt; 6&lt;br&gt;3 = 4-10 mm, n ≤ 5&lt;br&gt;4 = 4 mm, n &gt; 6&lt;br&gt;5 = &gt; 11 mm, n &gt;1&lt;br&gt;6 = confluent</td>
</tr>
<tr>
<td>Temporal</td>
<td>0 = absent&lt;br&gt;1 = &lt; 3 mm, n ≤ 5&lt;br&gt;2 = &lt;3 mm, n &gt; 6&lt;br&gt;3 = 4-10 mm, n ≤ 5&lt;br&gt;4 = 4 mm, n &gt; 6&lt;br&gt;5 = &gt; 11 mm, n &gt;1&lt;br&gt;6 = confluent</td>
</tr>
<tr>
<td>Internal capsule</td>
<td>0 = absent&lt;br&gt;1 = &lt; 3 mm, n ≤ 5&lt;br&gt;2 = &lt;3 mm, n &gt; 6&lt;br&gt;3 = 4-10 mm, n ≤ 5&lt;br&gt;4 = 4 mm, n &gt; 6&lt;br&gt;5 = &gt; 11 mm, n &gt;1&lt;br&gt;6 = confluent</td>
</tr>
</tbody>
</table>

*Figure 17.* Example of Scheltens ratings for white matter hyperintensities in pediatric TBI participants. Note that sample images are only a representation of WMHs as a single image is often unable to capture all WMHs located in the brain; n = number of hyperintensities; Bolded text shows the scoring by rater.
Appendix F: Log Transformed Correlation between Manual Tracing and NeuroQuant®

Figure 18. Log transformed correlation between manual tracing and NeuroQuant® WMH volumes
Appendix G: Summary of Neuroimaging Data with Outliers Excluded

Extreme outliers were observed on measures of WMH, including Scheltens (skewness and Kurtosis; 3.66, 17.66), NeuroQuant® WMH (2.85, 8.29), and manual tracing (6.36, 42.44). Figure 9 shows outliers >3 standard deviations (SD) away by method. As statistical analyses can be influenced by outliers, all analyses were run with outliers excluded for statistical thoroughness. Note that participants with >3 SD of WMH often had the largest amount of WM pathology and their findings may reflect more severe brain injury.
Figure 19. Scatterplot of WM methods by injury severity (GCS). Participant A had outlying data for both Scheltens and manual tracing, participants (B, C, & D) had outlying NeuroQuant® data. Two participants (B & C) had the same GCS score and outlying NeuroQuant® WMH volumes as shown by the connected blue arrow. Highlighted areas on MRI scans indicate WMHs as identified by their particular method (Ex: manual, NeuroQuant®, or Scheltens); WMH: white matter hyperintensity; 3SD: three standard deviations.

Table 10

Volumetric Data for Each Method with Outliers Excluded, Grouped by Injury Severity and Total

<table>
<thead>
<tr>
<th></th>
<th>Complicated-Mild M (SD)</th>
<th>Moderate-Severe M (SD)</th>
<th>Total M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheltens Ratings</td>
<td>1.90 (1.81)</td>
<td>4.33 (3.81)</td>
<td>3.03 (3.14)</td>
</tr>
<tr>
<td>n = 31</td>
<td></td>
<td>n = 27</td>
<td>n = 58</td>
</tr>
<tr>
<td>Manual- WMH Volumes</td>
<td>.01 (.01)</td>
<td>.03 (.09)</td>
<td>.02 (.06)</td>
</tr>
<tr>
<td>n = 30</td>
<td></td>
<td>n = 26</td>
<td>n = 56</td>
</tr>
<tr>
<td>NeuroQuant®-WMH</td>
<td>.06 (.05)</td>
<td>.06 (.03)</td>
<td>.06 (.04)</td>
</tr>
<tr>
<td>Volumes</td>
<td>n = 30</td>
<td>n = 25</td>
<td>n = 55</td>
</tr>
<tr>
<td>NeuroQuant®- Total</td>
<td>27.83 (2.42)</td>
<td>28.68 (2.47)</td>
<td>28.10 (2.45)</td>
</tr>
<tr>
<td>WM Volumes</td>
<td>n = 31</td>
<td>n = 9</td>
<td>n = 58</td>
</tr>
<tr>
<td>FreeSurfer- Total WM</td>
<td>29.04 (1.43)</td>
<td>28.37 (1.30)</td>
<td>28.51 (1.72)</td>
</tr>
<tr>
<td>Volumes</td>
<td>n = 32</td>
<td>n = 9</td>
<td>n = 59</td>
</tr>
</tbody>
</table>

Notes. Total WM volumes did not change as there were no outliers to exclude; WMH = white matter hyperintensities; WM = white matter; M = mean; SD = standard deviation; n = sample size; volumes are cm³
Figure 20. Scatterplots with regression line for each WMH comparison with outliers excluded A)

Scatterplot with regression line show positive correlation between manually traced and NeuroQuant® B) Scatterplot of Scheltens ratings and manual WMH volumes with regression
line shows a positive correlation; C) Scatterplot of Scheltens ratings and NeuroQuant® WMH volumes with regression line shows a positive correlation; WMH: white matter hyperintensities

Table 11

Correlation and Inter Class Correlations (ICC) with Transformations with Outliers Removed

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Correlation</th>
<th>Log</th>
<th>z(log)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ICC (n)</td>
<td>ICC (n)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>Manual – NeuroQuant®</td>
<td>n = 54, r = .12</td>
<td>ICC (38) = .11</td>
<td>ICC (n) =</td>
</tr>
<tr>
<td></td>
<td>p = .23</td>
<td>CI = -1.59 to .39</td>
<td>(n) =</td>
</tr>
<tr>
<td>Manual – Scheltens</td>
<td>n = 56, r = .48</td>
<td>ICC (36) = .75</td>
<td>CI = .51, .87</td>
</tr>
<tr>
<td></td>
<td>p = .00</td>
<td>(Excellent)</td>
<td>(Excellent)</td>
</tr>
<tr>
<td>NeuroQuant® – Scheltens</td>
<td>n = 54, r = .23</td>
<td>ICC (47) = .42</td>
<td>CI = -.04 to .68</td>
</tr>
<tr>
<td></td>
<td>p = .03</td>
<td>(Fair)</td>
<td>(Fair)</td>
</tr>
</tbody>
</table>

Note: There are no changes in the interpretation of ICC agreement between the analysis run with or without outliers; z-log = z-scores of the log transformed data.