2020

Neck Muscle Composition in Persistent Whiplash Associated Disorder: A Relationship with Disability

Bradford G. Callan
Nova Southeastern University

Neck Muscle Composition in Persistent Whiplash Associated Disorder: A Relationship with Disability

By:
Bradford G. Callan

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Nova Southeastern University
Dr. Pallavi Patel College of Health Care Sciences
Physical Therapy Department

April 2020

Dr. Pallavi Patel College of Health Care Sciences
Department of Physical Therapy

We hereby certify that this dissertation, submitted by Brad Callan, conforms to acceptable standards and is fully adequate in scope and quality to fulfill the dissertation requirement for the degree of Doctor of Philosophy in Physical Therapy.

_____ _____
Dr. Joshua Cleland, PT, PhD, FAPTA — Date
Chairperson of Dissertation Committee

_____ _____
Dr. Morey J. Kolber, PT, PhD, OCS — Date
Member, Dissertation Committee

_____ _____
Dr. James Elliot, PT, PhD, FAPTA — Date
Member, Dissertation Committee

_____ _____
Dr. Dave Walton, PT, MSc, PhD — Date
Member, Dissertation Committee

Approved:

_____ _____
Dr. M. Samuel Cheng, PT, MS, ScD — Date
Director, Physical Therapy PhD Program

_____ _____
Dr. Shari Rone-Adams, PT, MHSA, DBA — Date
Chair, Department of Physical Therapy

_____ _____
Dr. Stanley H. Wilson, PT, EdD, CEAS — Date
Dean and Associate Professor

2020

Abstract

Title: Neck Muscle Composition in Persistent Whiplash Associated Disorder: Relationship with Neck-Related Disability

Purpose/hypothesis: Neck pain and related symptoms following a motor vehicle collision (MVC) can significantly influence the quality of life for some people. While the MVC related mechanics of the head and neck gives rise to the term whiplash, the signs and symptoms of neck related disability are often complex and are clinically referred to as whiplash associated disorder (WAD). Muscle fatty infiltrate (MFI) has been associated with persistent WAD, but its influence on the generation and maintenance of WAD is largely unknown. The purpose of this study was to evaluate the relationship between MFI and baseline demographic variables related to persistent WAD.

Methods: 97 Participants presenting to an academic emergency medicine department in Chicago, IL enrolled in a parent longitudinal study investigating recovery from whiplash injury (ClinicalTrials.gov Identifier: NCT02157038). Within 1-week of the MVC, an MRI of the cervical spine was performed to quantify the percentage of MFI. Baseline demographics included: Neck Disability Index (NDI) scores, numeric pain rating scale (NPRS), age, sex, sleep disturbance scores, and BMI. At 2-weeks, 3 and 12-months post-MVC, NDI and NPRS scores were collected along with a co-registered MRI of the cervical spine. Final group membership was based on 12-month NDI scores: Recovered ($< 10\%$) or persistent WAD ($\geq 10\%$). Using logistic regression, at each time point, the variables were evaluated to determine significance associated with persistent WAD along with effect sizes.

Results: At baseline, variables found to be predictive for the persistent WAD group were: Female sex, increased MFI, and sleep disturbance. For persistent group classification, an R-squared of 0.27 for the model was reported as was odds ratios (OR) for individual variables ranging from 1.1 to 4.38. Across the 12 months, except for sleep at 2 weeks, the effects of MFI, sex, and sleep disturbance were significant, and R-squared increased to 0.56.

Conclusion: Results demonstrate an association with females, increased MFI, and sleep disturbance on persistent WAD. The variables maintained their significance across time along with increasing R-squared values related to the prediction of persistent WAD.

Acknowledgements

I would like to take the time to state that completing this PhD did not occur in isolation, it takes a village to accomplish a task such as this one, and I firmly believe that I would not have been able to complete all of the required work without the assistance and support of my village.

First, thank you to my committee Dr. Josh Cleland, Dr. James Elliott, Dr. Dave Walton, and Dr. Morey Kolber. Dr. Cleland your steady hand helped to keep the ship going straight and on time. Dr. Elliott's expertise on subject matter and written revisions was invaluable. Dr. Walton provided much needed assistance with statistical analysis along with lighthearted comments to keep an upbeat mood. Dr. Kolber and his intimate knowledge of the Nova process allowed me to manage the path without any significant setbacks. It is with your guidance, expertise, and encouragement that I was able to work my way through many of the difficulties that arose during this project. You knew when to help me and when to let me figure it out for myself, and for that I will be forever grateful.

Beyond my formal committee members, my friends and colleagues at Central Washington University were so very giving of their time in aiding me. Dr. Dominic Klyve reviewed much of my statistical analysis and Dr. Loretta Gray assisted in editing my writing, both of whom were instrumental in helping me. As well, all of my co-workers and friends who have listened to me bounce ideas around as I worked through problems, please know I am so very thankful for your time and support. A special thank you to Dr. David Hosmer at the University of Massachusetts, it was priceless to have assistance from a man who literally wrote the book on logistics regression.

To the friends I have gained over the years at Nova, I feel blessed that I was able to go through this process with you. I was so lucky to have met all of you, and I cannot imagine a better group of people to work through this program as we were able to call on each other at any time to provide support and assistance; I look forward to the years ahead as our careers move onwards.

I want to thank my mom and dad who taught me the value of education and the desire to push myself and work towards my goals. Finally, thank you to my daughters, Jenna and Ashley, for their support over the years. I could not have accomplished this without your love and understanding, there were difficult times and days when you two were the only thing that kept me going. Thank you for being you, I would not want you any other way.

Table of Contents

List of Tables

List of Table continued:

List of Figures

Chapter 1: Introduction and Problem Statement

1.1 Introduction

Following a motor vehicle collision (MVC), occupants of the vehicles often report a wide spectrum of problems related to their neck, such as pain, range of motion loss, and tenderness.[1,2] These issues, in addition to other signs and symptoms, including headaches, dizziness, unsteadiness, cognitive difficulties, insomnia, depression, fatigue, anxiety, weakness, and other neurological deficits are collectively known as whiplash associated disorder (WAD).[2,3] Of those diagnosed with WAD, approximately 50% are expected to fully recovery within 3 months of the MVC, while the other 50% experience persistent pain and disability[4] with few diagnostic or theranostic options.

While evaluating WAD, imaging has not provided consistent information regarding the diagnosis and prognosis of WAD.[5,6] The lack of findings may be attributed to the variety of measures and protocols that have been used to collect the information.[7] Radiographs, although beneficial in showing gross fractures, are not sensitive enough to detect subtle lesions that may be present following a whiplash trauma.[8] However, due to high levels of ionizing radiation, computerized tomography (CT) is not appropriate for serial use.[9] On the other hand, magnetic resonance imaging (MRI) allows for greater clarity and tissue resolution,[10] but structural pathology is often visible even in people that are asymptomatic at the time of imaging,[11] thus it can be difficult to determine what is clinically meaningful.

A unique perspective with emerging evidence in imaging of the neck has revealed elevated levels of muscle fatty infiltrate (MFI) in those with persistent WAD,[12] and these levels

of MFI are related to increased self-reported disability as measured with the Neck Disability Index (NDI).[13]

In trying to understand WAD, having knowledge about individual variables is beneficial, but the next phase in the research process is to appreciate the interactions between variables, how these interactions can change over time, and if these variables can be altered to change a person's long-term prognosis. This dissertation seeks to gain a deeper understanding of the relationship between MFI, self-reported disability, the influence of demographic and psychosocial factors, and the changing importance of these relationships over time.

1.2 Problem Statement

WAD is a complex biopsychosocial (BPS) phenomenon that is difficult to evaluate and treat. Research does not often consider the heterogeneity of the condition and is regularly focused on a single biological,[14] psychological,[15] and to a lesser extent socioenvironmental factor.[16] Using this model, many single variables have been found to be prognostic of poor recovery with WAD, but they are often of little value to the patient or clinician due to the variable being unmodifiable, such as age or sex, or they explain only a small amount of the variance in the outcome.[17] A more comprehensive understanding of BPS factors and their interactions is needed to understand the recovery pathways and to have the potential to be modified to assist in treatment planning.

From a parent investigation with an existing longitudinal dataset, this study consisted of a heterogenous population of 97 participants with varying levels of WAD related disability. With data collected from these participants, baseline measurements were used to determine what, if any, relationship exists between MFI, disability, and BPS factors. With this information, the

relationships were examined at baseline and at repeated points over a 12-month period following the MVC to more fully understand their influence on WAD symptoms.

1.2.1 Relevance

MFI has been found to have a significant relationship with the transition to chronicity in those with WAD,[18,19] as well, increased rates of MFI have been found to be correlated with elevated NDI scores.[19] This study will address gaps in the literature related to MFI by understanding potential interactions with psychosocial and sociodemographic factors that have been found to have relationships with the development of chronic WAD symptoms. The study evaluated the relationships not just at a single time point, instead, the longitudinal associations of these relationships were explored to understand the time-varying differences in magnitude of these associations. By studying these links, we obtained a greater understanding of WAD and its progression within a person rather than the isolated variables that have been previously found and used to describe the presentation of WAD.

1.2.2 Research questions

1) Evaluate the combined relationship between the baseline quantity of MFI and various BPS factors on NDI scores.

2) Assess if the relationships between MFI, BPS factors, and NDI scores changes in magnitude at specific time points over a 12-month period.

1.2.3 Hypotheses

MFI has been found to have a relationship with the transition from acute to chronic WAD symptoms,[19,20] as well, many other variables have been found to be associated with its development.[17] While these factors have been examined individually, combing factors to create a

more robust view of WAD has not been adequately examined. Yet at this time, there is no evidence to support the idea that combining factors will generate a superior understanding of WAD. Therefore, we hypothesize that the interactions between MFI and the BPS factors will produce a weak to non-significant relationship with NDI scores at baseline and at the multiple time periods over the 12-months following the MVC.

1.3 Whiplash and Whiplash Associated Disorder

The term whiplash is used to describe the transfer of energy from the thorax through the cervical spine resulting in a rapid acceleration and deceleration of the head and neck.[2,21] The most frequent application of this force occurs during an MVC, typically in a rear-end collision, but it can also occur during a side-strike MVC, sporting events, and after a fall.[2]

Following a rear-end MVC, 40-67% of occupants of the impacted vehicle will demonstrate no signs or symptoms of injury.[22,23] However, if someone does develop signs and symptoms about the head and neck following an MVC they are most commonly diagnosed with WAD. The diagnosis is based upon the clinical presentation of the patient and not on any standardized tests or measures as there are currently no objective gold-standards available to definitively diagnosis WAD.

1.3.1 Tissue Trauma and Imaging

The role of tissue trauma in the cervical spine following an MVC and its relationship to WAD symptoms is a research line that has produced contradictory outcomes with various studies. Imaging studies routinely do not find evidence of structural trauma following an MVC,[5,24,25] but invasive studies, cadaver or surgical based, following an MVC demonstrate numerous non-lethal structural injuries to bones, facet joints, joint capsules, ligaments, discs,

cartilage, muscles, nerves, and vascular structures.[8,26] As well, other studies have shown significant relief in WAD symptoms following injections or nerve ablations in the neck[6,27] which imply that damage to the cervical tissue is responsible for at least some of the symptoms . Although these injuries have been visualized in surgery and influenced with injections, there is still great debate about their role in chronic WAD symptoms.

Numerous studies performed over the years have found weak to non-existent associations between findings on imaging and self-reported pain, disability scores, and various signs and symptoms that people ascribe to the MVC.[5,24,25] However, these studies were often exploring a single lesion to try and explain the symptoms of WAD. For example, previous author groups have explored the alar ligament and cervical facet joints following an MVC for their ability to reproduce symptoms related to WAD.[14,28] These studies have not traditionally accounted for psychosocial factors that can alter the pain experience such as PTSD or pain catastrophizing.[17,29] As well, prior studies did not commonly consider the pre-MVC status of the subjects and how these factors can alter the long-term prognosis.[30] By not understanding how the pre-injury tissue, genetics, and psychological makeup of the person can alter the WAD experience, the research outcomes may be incomplete and it is understandable why imaging studies have not been beneficial in explaining WAD symptoms.

The MRI as an imaging modality is the safest choice for patients and provides the greatest clarity for evaluating the cervical spine following a whiplash injury in both the clinical and research settings.[10] However, there are many variables and parameters within the machine that can help to create an optimal image. Some of the parameters used to generate images are slice thickness, signal-to-noise ratio (SNR), total scan time, the types and positioning of coils used to receive the signal, strength of the magnetic field, as well as the computational power of

the processing computers and the applicable software.[10] With these parameters, an alteration in one will usually result in a change to another which forces researchers to carefully consider the most optimal balance between the variables.[10] The numerous variations available with the parameters demonstrate that there is no perfect scanning sequence for everyone. Instead, all the parameters must be considered and modified to determine the most beneficial scan sequence for the given problem at hand.

Although the MRI can provide great clarity, evaluating imaging research can be difficult due to the use of variable protocols during scans. The variability can create difficulties with comparing the findings of several studies such as with a meta-analysis.[5,24] Therefore, a minimum set of standards has been proposed to be reported to allow for more direct comparisons between studies.[31] Within this dissertation, the collected data and images followed a rigorous standardized protocol to maximize image clarity as well as allow for co-registration.

1.3.2 Muscle Fatty Infiltrate

MFI is a process where adipose tissue is deposited below fascia, and between muscle fibers and muscle groups.[32] MFI is a naturally occurring physiological attribute that is present in all people and is associated with cellular metabolism, maintaining physiological homeostasis, and may be associated with the inflammatory cascade following tissue trauma.[32] By depositing fatty tissue within the muscle, the muscle contractility may be modified which can alter biomechanics,[19] and in turn can create abnormal pulling, shearing, and compression throughout the tissues and may result in some of the signs and symptoms associated with WAD.[33]

It is unknown specifically why MFI develops, and there is even less understanding of why it develops in greater proportions in people diagnosed with persistent WAD.[12] There are numerous theories for why MFI develops including disuse atrophy,[34] reactions to the

inflammatory processes,[32,35] or it may be mediated by a combination of BPS variables and the stress response system.[29]

The exact role of MFI and the maintenance of WAD is not well understood. In studies by Elliott and colleagues, those with persistent WAD demonstrated MFI in significantly greater amounts than those with insidious onset of neck pain as well as having a relationship with elevated NDI scores.[12,18] This discovery promotes the idea that a mechanism related to the trauma may initiate the increased deposits of MFI and that it can be a factor to the prolonged symptoms associated with WAD.[20]

Beyond the decreased use of the neck muscles due to pain resulting in disuse atrophy, MFI is theorized to result from several potential mechanisms. The first is an altered stress-response system due to a distorted activation of the hypothalamic-pituitary-adrenal (HPA) axis which can alter the sensory response for those with a whiplash injury.[36] The altered HPA axis and sensory response can result in increased pain with movement, decreased activity, and atrophy of the muscles with deposits of MFI as the final outcome.

MFI can also result from negative or maladaptive thoughts, attitudes, and actions about pain and movement that lead to avoidance of activity.[29] There are several theories that describe this topic including the pain adaptation model,[37] the fear avoidance (FA) model,[38] and the pain avoidance model.[39] All three of these models have subtle nuances that separate each one from the others, but in general they describe how maladaptive thoughts and beliefs of pain can alter the willingness to move. By demonstrating a decreased willingness to move, patients decrease their level of activity, which can lead to disuse atrophy, weakness, more difficulty with movement, and in the end more pain with movement. Within all the models, patients demonstrate some level

of fear and anxiety related to movement, and this fear and anxiety can alter the stress response system creating a physiological response which can alter their pain threshold.[40]

Finally, the role of the sympathetic nervous system (SNS) in the development of MFI must be considered.[41] The SNS has a primary role of activating the fight or flight system when a person is in a stressful situation.[40] Numerous stressors are known to alter the SNS, but for those that are afflicted with WAD, the most commonly elevated stressors when compared to a non-injured control group are anxiety, depression, pain catastrophizing, and PTSD.[29] It is theorized that MFI and its relationship to WAD may be mediated by PTSD,[19] but it most likely has complex interactions with several BPS variables which can be combined to form a more robust picture of the patient.

1.3.3 Biopsychosocial Factors

With an MVC, despite the similar rear-end collision mechanism, the signs and symptoms associated with WAD are not identical for each person. If the signs and symptoms were based upon a single lesion, then a targeted intervention should consistently improve recovery. Evaluating WAD with this singular view, the current state of research provides varying levels of support for or against numerous treatments including immobilization using a collar,[42] manipulations of the cervical or thoracic spine,[43] stabilization exercises,[44] reassuring advice to "act as usual,"[43] and injections.[28]

Albeit limited, most all treatments have had some level of improvement for the patient, but no treatment has been consistently successful. This idea is reflected in the population-based studies that demonstrate the rates of recovery for those diagnosed with WAD have not significantly improved in 30 years.[4,45] It can be argued that the lack of improvement in outcomes lends support to the idea that WAD is not a single homogenous entity. Instead, a more inclusive

view of the patient must be obtained that takes into consideration all potential influencing domains to allow for proper evaluation and clinical management.

The term BPS represents the biological, psychological, and social factors that are part of an individual. All three areas can affect the WAD experience, therefore, they all must be examined to understand the complexities and heterogeneity of the chronic WAD presentation. Although each area can be evaluated on its own merits, in the end, there is an understanding that all the variables likely interact and come together to form a complete clinical picture that is unique for each person. Currently, it is unknown which interactions exist and are relevant, and if they are static or fluctuate over time. These interactions may help create the overall makeup of the person which may be considered greater than the sum of the individual parts. To generate the most relevant information regarding the understanding of WAD, this dissertation discovered how various BPS factors interact, specifically the role of MFI and sociodemographic variables, and how they change over time to influence the WAD experience.

1.4 Summary

The overall understanding of WAD has transformed in recent years with the most prominent change being the acceptance of the role of psychosocial variables in the manifestation of a person's symptoms. For years, when these variables were described in regards to WAD, it was felt that the person was making-up their symptoms as they were either malingering, hysterical, faking it for monetary compensation, or to gain attention.[46] However, it has been shown that the psychosocial factors can effect cognitive and neurobiological processes, including the physiological responses to painful stimulation.[40] No single variable explains a significant amount of the variance associated with WAD, but some factors may be more relevant than others

in explaining the overall experience. However, currently most of this evidence is correlational with very little support for causal associations.

Beyond the psychosocial factors related to WAD is the limited understanding of how the biological factors can play a part in the WAD experience, specifically the role of MFI and its evolving relationship to symptoms. It has been discovered that MFI has a relationship with the transition to chronicity, but how MFI interacts with other known psychosocial variables and how these relationships change over time is unknown.

Therefore, instead of focusing on a single BPS factor as the cause of a patients' symptoms, diagnosis and management of those with WAD will need to consider the potentially fluid relationships between numerous variables for each person. Thus, WAD cannot be viewed as a single construct that is identical in each person, instead those with WAD should be viewed as a heterogenous population whose symptoms will manifest uniquely based on their own personal experiences.

1.5 Definition of Terms

Acute – Symptoms that are present between 3 and 4 weeks following an initial injury.

Anxiety – A mental health condition that manifests in an atypical emotional response from the anticipation of a future threat.[47]

Artifacts – Something observed that is not naturally there, but instead is present due to a distortion in the image, hardware/software malfunction, or environmental influences.[10]

Autonomic nervous system (ANS) – Portion of the nervous system that manages bodily functions that are not under conscious control such as breathing and digestion.

Biopsychosocial (BPS) – An interaction of biological, psychological, and social factors that work individually or together to alter the outcome of an injury or disease process.

Central nervous system (CNS) – Consists of the brain and spinal cord and controls most bodily functions including, but not limited to, thoughts, movements, sensations, and body awareness.

Chronic – Symptoms that persist for more than 3 months.

Co-registration – Evaluating serial images on a single grid to determine if changes occur over time.

Depression – A mental health condition that results in sadness and accompanies somatic and cognitive changes that affect the person's ability to function.[47]

Fear avoidance beliefs – An exaggerated or fearful reaction to pain perception in which the person has an extreme fear of pain related to movement, so much so that they are unwilling to move, and they avoid activities that may exacerbate their symptoms.[38]

Muscle fatty infiltration (MFI) – A process where fat is deposited below the fascia, and between muscle fibers and muscle groups.[32]

Magnetic resonance imaging (MRI) – An imaging tool that uses magnets and radiofrequency pulses to generate a digital picture of the internal structures of the human body.[10]

Motor vehicle collision (MVC) – Occurs when two vehicles strike one another, and the collision may or may not cause damage to the vehicles.

Neck Disability Index (NDI) – A self-report outcome measure for evaluating the functional ability for those with neck pain.[48]

Pain catastrophizing – The belief system that this pain will likely lead to a lifetime of disability, and if something is painful that it must be causing damage or making the injury worse, therefore the pain should be avoided.[49]

Peripheral nervous system (PNS) – The portion of the nervous system that is outside of the central nervous system (brain and spinal cord).

Post-Traumatic Stress Diagnostic Scale (PDS) – A self-report scale measuring the symptoms associated with PTSD.[50]

Post-Traumatic Stress Disorder (PTSD) – A psychological condition with persistent mental and emotional stress because of an injury or severe shock resulting in a constant recall of the traumatic experience.[47]

Rapid eye movement (REM) – A period with a sleep cycle that is theorized to be important for memory consolidation as well as helping to maintain homeostasis across various systems.[51]

Slow wave sleep (SWS) – A distinct portion of the sleep cycle which physically contrasts with REM cycle, but also is theorized to assist with maintaining the bodies normal physiological activities.[51]

Sympathetic nervous system (SNS) – One portion of the autonomic nervous system that alters numerous hormones and physiological responses in response to stressful events which results in activating the fight or flight phenomenon.

Whiplash Associated Disorder (WAD) – A group of signs and symptoms that a person will present with following a whiplash event. The signs and symptoms can include, pain, stiffness, decreased range of motion, tenderness, as well as headaches, dizziness, and weakness.[2]

Whiplash – A transfer of energy from the thorax to the cervical spine and head through a rapid acceleration and deceleration.[2]

Chapter 2: Review of the Literature

2.1 Introduction

Whiplash and its associated disorders (WAD) represent a complex biopsychosocial (BPS) condition commonly resulting from a motor vehicle collision (MVC). In the United States (US), it is estimated between 1 million to over 4 million people a year will be injured in a MVC,[52-54] of which only 50% of those with WAD recover within the first 6-12 weeks.[4] The other 50% will continue to have persistent symptoms at least 12 months after the collision.[4] Of those with persistent symptoms, up to 25% will have a more severe clinical presentation in the long term[4] as well as significant medical and legal costs and lost income due to time away from work.[55,56] It is estimated that these costs collectively can range between 10 and 29 billion US dollars per annum.[55,56] Lessening the transition from acute to chronic WAD is of the utmost importance to decrease the personal, societal, and economic burden related to these injuries

While the mechanism of bodily injury during a MVC are fairly well understood,[57-59] it is unknown why 25% of people who report WAD symptoms following an MVC report persistent, severe symptoms. Trauma to a number of vulnerable tissues such as ligaments, discs, and joint capsules is speculated to be a mechanism for the pain and loss of function following an MVC.[8,24,60] A vehicle collision can create easily identifiable injuries such as lacerations or fractures,[61] but, occult pathoanatomical injuries to the vertebral disc or zygapophyseal (z) joints may exist and have clinical management implications.[62] However, a lack of identification with a clinical examination or diagnostic imaging has led to skepticism of any consistent structural lesions.[24,63]

Beyond the potential physical injuries, numerous psychological and social factors have been studied in an attempt to understand their influence on WAD symptoms.[17,64] The presence of psychological factors, such as post-traumatic stress disorder (PTSD) and pain catastrophizing,[17] have temporally been associated with the subsequent development of chronic WAD symptoms.[40,64,65] However, recognizing the relationships between these variables and WAD has not improved management strategies to reduce the rates of chronic symptoms.[4,45] Whether it be physical or psychosocial factors, no two situations are exactly alike; thus, finding a single mechanism to explain WAD is unlikely. The lack of an easily identified single factor may explain our limited understanding of how to assess and manage these cases on a patient-by-patient basis.

The purpose of this chapter is to provide a brief background on the history of WAD, including what is currently known about the condition. A description and review of the clinical features associated with acute WAD and the transition to chronicity including BPS factors, imaging findings, and the potential physical, physiological, and psychological changes will be examined. Lastly, a discussion on the importance of combining BPS variables to create a more comprehensive picture of WAD will be presented.

2.2 History of WAD

Within this historical overview, a comparison will be made between the modern day whiplash injury and its associated signs and symptoms with the description of "railway spine" put forth in the 19th century by the British surgeon J.E. Erichsen.[46,66] What is of interest within the historical understanding is that although these injuries have been described for well-over 100 years, the overall knowledge and management of WAD continues to perplex both clinicians, patients, and other stakeholders.

Much like the constellation of symptoms seen with WAD, Erichsen attributed the similar signs and symptoms (neck pain, stiffness, tenderness, and fatigue) to the "accident" on the railway.[46,66] Individuals diagnosed with railway spine often lacked any visible injury that could explain their complaints, thus Erichsen provided a vague description of what happened to them by stating "The whole system received a severe shake or shock" which resulted in "molecular disarrangement."[42,66] It was felt this mechanism was responsible for the variety of reported signs and symptoms experienced after the rail accident. Although Erichsen lacked the technological advancements of today, his ideas of trauma to the entire system were similar to the current theories regarding the heterogeneity of those with WAD.[17,30,67]

It was not until a medical conference in 1928 when the term "whiplash" was used by H.E. Crowe to describe the mechanism of injury (MOI) as a rapid movement of the head and neck that occurs during an MVC.[46] However, the term was not formally used in the medical literature until 1945,[66] but since then, the meaning has evolved as the term whiplash is currently used to describe not only the MOI, but also the injury itself.

Using the term whiplash as a diagnosis over the years has been problematic, in part because there is no gold standard for determining who has the injury. Currently, the diagnosis is primarily based on two factors: 1) an association with a trauma affecting the neck, and 2) the responses provided by the patient, including reports of neck pain, tenderness, or decreased range of motion.[2] However, WAD can result in various other signs and symptoms, including fatigue, insomnia, or difficulty concentrating;[3] thus, basing the diagnosis on the presence of neck pain alone may be inadequate.

The difficulties in identifying those with WAD has persisted despite advancements in imaging and research designs which were assumed to be of benefit in determining relationships

between trauma and symptoms. As imaging techniques improved, a plausible idea was postulated by researchers that a single lesion would be discovered that could explain the presenting signs and symptoms. For example, Krakenes reported that damage to the upper cervical ligaments, in particular the alar ligament, were correlated with chronic WAD symptoms and would "contribute" to the patient's presentation.[14] However, subsequent imaging studies have not corroborated these results nor discovered any other structural lesion that can account for the presentation of chronic WAD,[5,8,24,25] and this lack of information has fueled the skepticism of WAD as a true injury.[63]

According to the most current research and theories, the signs and symptoms associated with WAD have been speculated to arise from physical, physiological, psychological, and social factors.[28,40,68,69] Multiple variables within all these domains have been found to be associated with WAD symptoms, but identifying a single cause and effect mechanism has been unsuccessful. The focus on a single mechanism may have unintentionally limited the evaluation of those with WAD, and the ability to clearly identify and understand the interactions of known BPS factors.

With these known limitations, the core focus of this dissertation was to examine several of the primary factors regarding the whiplash conundrum. Through a robust data set, the role of muscle fatty infiltrate (MFI) and its associations with psychosocial factors were examined while attempting to provide a multifactorial picture of WAD during the transition to chronicity. With this information, appreciation of the heterogeneity of those afflicted with WAD may be more fully developed and can contribute to its overall understanding.

2.3 Clinical Features and Classifications of Persistent WAD

2.3.1 Mechanism of Injury

The mechanics of a rear-end MVC have been analyzed using high speed cameras with human subjects, cadavers, animals, human analogs, and computer simulations.[8,26,62,70] During a rear-end MVC the head moves not only in the directions of extension and flexion, but the torso is thrust upwards relative to the body which results in a buckling of the cervical spine and the generation of a non-physiological S-shaped curve.[57,71] The S-shaped curve has the potential, via shear and compression forces, to create abnormal stresses within the ligaments, discs, bones, and joint capsules.[59,72] Individual forces are not necessarily going to cause an injury, as the discs and facet joints of the neck are designed to resist some compression and shear.[73] Instead, it is the combination of the shear and compression forces applied through a non-physiological curvature of the neck that has the potential to damage the tissue. If the tissue is damaged, the inflammatory cascade may be triggered resulting in increased nociceptive sensitivity and may result in pain for the person.[74]

2.3.2 Clinical Features of WAD

The controversy surrounding WAD is often based on the subjectivity of the clinical presentation. Other than neurological changes, such as altered reflexes or loss of strength, most of the clinical findings are self-reported, including pain, tenderness, headaches, dizziness, insomnia, cognitive difficulties, unsteadiness, depression, fatigue, and anxiety.[2,3] It is often difficult to say the collision caused these symptoms because many people who have no history of WAD or a whiplash injury present with these same complaints. With that in mind, due to the inability to visualize a specific lesion related to WAD, there is hesitation in accepting it as a legitimate injury.

To complicate the situation further, patients can often present with a variety of these symptoms, and it is uncommon to have two patients present identically. Therefore, instead of trying to understand WAD based on symptoms, it may be more beneficial to focus on known risk factors associated with chronic WAD. These risk factors and their association with prognosis may provide increased benefit for both clinicians and patients who are trying to understand the condition and decrease the rates of chronicity.

2.3.3 Classification of Persistent WAD

The Quebec Task Force (QTF) produced a monograph in 1995 with a goal of synthesizing the best available evidence and consensus opinions regarding risk, diagnosis, treatment, and prognosis of injuries sustained in a MVC, specifically related to whiplash trauma.[2] Of all the findings reported, the creation of the 0 to 4 grading scale used to describe WAD (Table 2.1) may be the most contentious.[2] The scale attempted to create a common language that could be spoken by all clinicians and researchers regarding the severity of the injuries sustained, but it is based primarily on self-reported signs and symptoms, such as neck pain, tenderness, decreased range of motion, or neurological changes. As the determining factors for the grades, it was implied the signs and symptoms were related to the trauma itself and the supposed tissue damage. In this grading schema there was no consideration given to potential influencing factors such as sleep disturbances, the physical or emotional health of the person, or psychological factors such as PTSD.

The QTF examined various prognostic variables, such as sociodemographic variables age, sex, income, and collision related data, including crash severity, direction of collision, and type of vehicles in the MVC.[2] Of these variables, most were found to be inconclusive regarding their influence on WAD symptoms, however, most of the research evaluated individual variables

and often only cited significance levels and did not report effect sizes.[2,17] Although the findings of the QTF were based on the research available at the time, recent data has found collision related variables have little to no effect on the prognosis of WAD.[17]

Table 2.1 Quebec Task Force Grading Scale For WAD[2]

Grade 0	No complaints about the neck; no physical signs
Grade 1	Neck complaint of pain, stiffness, or tenderness only; no physical signs present
Grade 2	Neck complaints AND musculoskeletal signs including decreased range of motion and point tenderness
Grade 3	Neck complaints AND neurological signs including decreased deep tendon reflexes, decreased sensation, or decreased strength
Grade 4	Neck complaints AND fracture or dislocation

Key: WAD – whiplash associated disorder

A primary goal of the QTF was to improve the management, outcomes, and rates of recovery for those with WAD.[2] But, the model of categorizing patients based on their symptoms has not produced the desired results given that the rates of transition to chronicity have not significantly improved since the grading scale was published.[4] A case can be made that it has some prognostic value, as there is evidence that the higher the grade of WAD the slower the recovery time,[75] but it does not provide any benefit for the management or outcomes for those with WAD.[76]

Moving beyond the QTF grading schemes, a clinical prediction rule (CPR) has been derived and validated for classifying a patients probability of transitioning to chronicity.[77,78] When all three criteria of the CPR are present, a positive predictive value of 91% has been reported;[78] this means that 91% of the people who scored positive on the CPR were in the non-recovery group. The sensitivity and specificity of the CPR have been reported as 44% and 99% respectively.[78] The strong value for specificity provides greater confidence in ruling someone in when they meet the CPR as the false positive rate is only 1%. However, the sensitivity is not

strong enough to confidently rule someone out of developing prolonged symptoms if they do not meet the CPR.

Within the CPR, the three primary predictors are a neck disability index (NDI) score of 40% or greater, 35 years of age or older, and a score of 6 or greater on the hyperarousal subscale of the post-traumatic stress diagnostic scale (PDS).[77] Although using three variables yields a strong positive predictive value, the CPR generates a high false negative rate of 56%, which means that 56% of the people who developed chronic WAD did not fulfill the CPR criteria. The high false negative is important as it means that people who develop chronic WAD are not being distinguished by the CPR and may not be receiving treatment. On the other hand, the high specificity results in a 1% false positive rate which means that these mis-identified people may end up receiving treatment when none is needed. In looking at the greater good, it seems that having a higher sensitivity (and a lower false negative rate) would be more beneficial as more people with the condition will be identified; thus, those that may benefit from treatment will receive it.

To improve the sensitivity of a test, the number of positive test results must be increased, and this can be accomplished in two ways: 1) lower the threshold for determining what is a positive result, or 2) use different factors. In lowering the threshold in an attempt minimize or eliminate false negatives, an extreme example would be to assume that all people in an MVC will develop WAD, this ensures all people have a positive result and a 100% sensitivity will be present. For a variety of reasons this is an unrealistic option, for the CPR, to decrease the positive threshold a solution could be to eliminate one or even two of the current factors within the model. However, these three factors were determined using a robust statistical analysis and it may be difficult to simply remove one or two of them.

If we cannot eliminate factors, perhaps using different factors that have not been fully explored may be of greater value in improving the sensitivity, such factors could include MFI or looking for interactions between BPS variables. Searching for these interactions is important because they have the potential to strengthen the relationships and prognostic ability more than an individual factor.[79] By not establishing potential interactions between the variables, it is possible the CPR may be incomplete and many people that transition to chronicity are not being identified, or treated, early in the process.

2.4 Current Understanding of WAD: Biopsychosocial and Imaging

2.4.1 Current Overview

After an MVC, a variety of both direct and indirect factors may be related to the signs and symptoms associated with WAD. Physical trauma, such as a torn joint capsule or ligament damage are some of the possible sequelae that can be attributed to the energy transfer between the motor vehicles and the people inside.[80] Although traumatic injuries to the neck have been visualized during surgery and in cadavers,[8] imaging has not been able to aid in diagnosing or understanding the relationship with WAD related symptoms.[5,24]

Beyond the initial trauma occurring from an MVC, secondary complications can arise, such as the increased levels of MFI, segmental instability, and it is theorized that post-traumatic arthritis can develop.[34,60,81] The complete understanding of how these issues can alter a person's pain experience and their functional abilities is not immediately recognized as the complications may take weeks or months to fully develop.

Outside of the physical injuries related to the MVC, psychological and social factors can also influence the symptoms associated with WAD.[17] Psychologically, PTSD, and to a lesser

degree various social factors, such as level of education, have all been found to be associated with symptoms.[17]

With these prognostic variables in mind, recovery trajectories have been estimated to determine who is more likely to progress to chronicity and who is more likely to recover without further incident.[4] Most variables demonstrate small and inconsistent effects when related to these trajectories, but two variables have been found to have a strong and consistent relationship with a poor prognosis: one, a high initial pain rating, 5.5 or greater out of 10, and two, an elevated NDI score, 14.5 or higher out of 50.[82] However, these variables in isolation offer limited usefulness in guiding treatment.

Although not as strong in their relationship to chronicity as the initial pain rating and an elevated NDI scores, PTSD, pain catastrophizing, and the development of MFI have a moderate association with chronicity, but they need to be understood more thoroughly to fully grasp their clinical relevance.[17] The underlying mechanisms of how these factors can influence WAD are theorized to be related to systems such as neurophysiological stress response,[40] altered motor control,[29] genetics,[83] and psychological maladaptive beliefs,[84] however the complete understanding is purely speculative at this time.

2.4.2 Muscle Fatty Infiltrate

MFI is a normal physiological attribute that is present in all people regardless of activity level, and it is recognized to have several primary functions including being a local energy reserve as well as an endocrine organ associate with maintaining homeostasis across multiple physiological systems.[32] It is also theorized to be part of the inflammatory cascade following a trauma by way of releasing proinflammatory cytokines, which when present for prolonged periods of time has the potential to alter the muscle tissue.[32] The degeneration in the muscle tissue

by way of increased MFI occur in two primary forms;[32] intra-myocellular where fat replaces muscle tissue, and extra-myocellular where fat is deposited between fascia or muscle fibers.[32] Regardless of its origin, MFI is likely to affect the quality of a muscle contraction which can result in diminished functional capacity (decreased strength, power, and endurance) for the person.[85-87] As well, in addition to altered or decreased muscle activation, the potential for abnormal mechanics could influence the long term health and function of the cervical spine.[88]

Sarcopenia is defined as a decrease in muscle mass and strength that often occurs as we age, and as it progresses, a loss of lean body mass and a concurrent increase in body fat and MFI is observed.[89] It is unknown what the exact biological mechanism is for sarcopenia.[89] However, it is generally recognized as a normal consequence of aging, although some view it as a disease, and the loss of muscle mass and strength, and increase in MFI, has been associated with declining scores on physical performance tests.[90] Thus, distinguishing between MFI that is present as a normal part of aging and that which is related to an MVC can be difficult, but recognizing the changes in functional scores as a result of MFI may explain some of the loss of function for those with WAD.

Following an injury, MFI can be deposited throughout the tissue and is theorized to be a result of an interaction between macrophages, various components of the inflammatory cascade, and fibro-adipocyte progenitors (FAP's).[91] The FAP's act similar to a stem cell and can influence the pathway in which an injured muscle fiber will either regenerate as a part of the contractile unit or differentiate into fatty tissue.[91] It is not known why FAP's may lead to increased levels of MFI following an injury, but it is plausible there may be genetic, epigenetic, or a combination of factors that predispose some people to the development of MFI following an MVC.

Disuse atrophy is another potential mechanism for altering the amount of MFI in tissue,[34] it is theorized that decreased use may result in myofibrillar atrophy and deposits of fat into the area.[92] However, this reasoning may be an overly simplistic view of the process as it does not account for the previously described factors (FAP's) that may influence MFI, and it does not explain the differences in MFI between two groups, one with WAD and the other with insidious onset neck pain.[12] It is presumed that the pain would result in similar rates of decreased use of their neck muscles and in theory equal amounts of disuse atrophy and MFI. However, the group with WAD demonstrated a significantly greater amount of MFI within the cervical musculature.[12] A second study reported that the amount of MFI in the cervical musculature was unrelated to the amount of range of motion lost following an MVC;[19] thus, the amount of disuse may be irrelevant to the development of MFI. With these observations, it is thought some mechanism of the trauma, the inflammatory cascade and interactions with FAP's, a neurophysiological response, genetics, an interaction of the stress-response system, an undiscovered factor, or any combination of these can influence the rate of MFI.[12,19,32,36]

Once MFI is present, it has the potential to propagate itself based on several potential mechanisms. The first is the release of proinflammatory cytokines that can promote a continued localized inflammatory response.[32] The local inflammatory response can have a catabolic effect on muscles which may result in increased MFI deposits,[93] pain,[94] and decreased mobility which can result in further atrophy of the muscles and potentially more MFI.[32]

Increased MFI following a trauma may also be related to the stress-response system and the effects of cortisol on the tissue.[36] Following an MVC, it is not uncommon for a person to experience transient or persistent distress regarding their physical or financial health, fear and anxiety regarding the collision, as well as various stressors present in their life prior to the MVC.

Any or all of these factors can alter the hypothalamus-pituitary-adrenal (HPA) axis which is the endocrine pathway that controls the release of cortisol.[40] Increased levels of cortisol have been reported following an MVC, and this hormone, similar to the proinflammatory cytokines, can alter the structure of muscle via its catabolic effects on the tissue thereby potentially resulting in increased deposits of MFI.[36,89]

Assessing the relationship between MFI and a functional outcome following an MVC has been observed. For example, all people who were in a recent MVC (less than 1 week) and evaluated with magnetic resonance imaging (MRI) of the cervical spine were found to have similar rates of MFI at baseline regardless of their long term outcomes.[20] However, by two weeks post MVC, a difference in the amount of MFI was observed between the three groups: recovered, mild, and moderate to severe symptoms.[20] With this information, a receiver operator characteristic (ROC) curve found a cut point of MFI at 2-weeks of 20.5% of cervical musculature to have a sensitivity of 87% for identifying the moderate to severe group, and a specificity of 92.9% for identification of the recovered group at 3-months.[20]

Currently, it is impossible to determine a cause and effect relationship between MFI and outcome measures, although an association between the two has been reported.[19,20] However, this relationship appears to be mediated by PTSD which means that the strength of the association between the two variables is stronger when these symptoms are accounted for.[20] The relationship between MFI, NDI, and PTSD is not understood, one theory is the quantity of MFI may be a result from the physical trauma of the MVC, and can result in PTSD along with increased pain and decreased function.[95] Conversely, the opposite may be true as higher levels of pain and NDI scores have been found to influence the amount of total fat within the muscles.[95]

The gold standard for assessing MFI is with a biopsy of the tissue, which is an invasive procedure, and currently, the best way to evaluate it conservatively is with MRI. However, with both methods, they must be assessed serially to determine if changes are occurring over time. Given the cost, time commitment, and observation that 50% of people will not develop long-term symptoms, it is unrealistic to believe all people involved in an MVC should obtain serial imaging or muscle biopsy as part of a routine standard of care.

Considering the relationship between MFI and chronic WAD symptoms, a clinical analog for MFI may prove important for identifying individuals at risk of developing long-term symptoms at a time when rehabilitation treatments may be more beneficial for the patient. Although preliminary, a small study of patients with WAD, who completed a 10-week exercise program, demonstrated significant changes in muscle structure with increased muscle cross-sectional area and decreased MFI in the cervical multifidus. The subjects also reported a subsequent decrease in NDI scores.[96]

Although MFI demonstrates significant prognostic value for people with WAD, it is only one component of the BPS profile. Understanding how MFI may be associated with other variables, such as age, sleep disturbances, or PTSD, may help develop a more complete clinical picture of WAD. However, it is unknown which variables will be the most beneficial for determining the prognosis of WAD, and it is not understood how variables may interact with each other or how the relationships evolve over time.

With that in mind, a focus of this dissertation is to evaluate how the interaction of several variables influence WAD, as well as how these interactions change over time. Ideally, a single measure would be provided soon after the MVC that could capture all domains and predict who will transition to chronicity which would allow treatment to be focused to that subgroup.

However, given the heterogeneity of patient presentation, pre-existing factors, and the changes that occur throughout time, (including neurophysiological, morphological, social, and psychological) it appears multiple measures will be needed to capture the evolving presentation of WAD. It could potentially be of benefit to understand which measures, when used together, create the most comprehensive and clinically meaningful assessment for those with WAD both cross-sectionally and longitudinally.

2.4.3 Biological and Psychological Factors

There are a numerous biological and psychological factor that have been reported to be associated with poorer outcomes for those with WAD. However, especially the biological variables, there have been none that have been noted to have consistent significant effects, as one report may cite an association while a follow up study finds no relationship. It is not fully understood why these varying results exist as it may have to do with populations within the studies, the determination of what it means to be recovered, or only looking at individual variables at single points in time. To fully understand the influence on WAD symptoms, several biological factors associated with MFI will be evaluated to fully elucidate the effect they have on outcomes.

An increase in a person's body mass index (BMI) has been associated with increased levels of MFI,[32] this may lead to morphological changes in the tissue and altered biomechanics and potentially prolonged symptoms that are present even before the MVC.[19] In addition to increased MFI, an elevated BMI may cause issues with persistent symptoms due to a larger habitus as well as systemic inflammation.[32] A larger habitus may limit joint or tissue mobility which can impede tissue healing, and elevated levels of systemic inflammation may maintain

symptoms due to persistent proinflammatory cytokines within the cells which can perpetuate a nociceptive pain response.[32]

Age is another biological factor that has been theorized and evaluated to determine its effect on WAD, and the results reported range from inconclusive, to no significant or small effects.[17] With age comes changes in our bodies including the previously described sarcopenia, disc degeneration,[97] decreased muscle mass,[98] and increased rates of chronic pain.[99] Based on conjecture at this time, all of these factors have the potential to create a symptomatic response following an MVC, but it is not certain if someone has one or all of these traits they will develop chronic WAD. Therefore, understanding how age and MFI may be related and associated with WAD may provide a greater understanding of long-term prognosis.

Psychological factors have long been evaluated and known to influence WAD symptoms, [42,46] but studies have tended to only evaluate single psychological variables to understand WAD and its various symptoms.[100] A few studies have attempted to address multiple variables with regard to prognosis, for example, Gargan combined neck range of motion and the scores on the General Health Questionnaire to help determine chronicity.[101] It was reported that the scores were much more beneficial at predicting the recovered group, 94% accuracy, than the group with chronic symptoms, 56% accuracy;[101] thus, it was not very useful clinically. A second study using a multivariate analysis reported several variables to be predicative of long-term symptoms.[102] However, a number of these variables, direction of impact and position of the occupant in the vehicle, have been found to not be associated with the prediction of symptoms.[17] As well, the measures of psychological distress were based on generic health outcome measures rather than item specific psychological measures, such as PTSD, thus their results are questionable.

Psychological issues such as depression and anxiety have been studied extensively in relation to WAD, but their presence is generally inconclusive in predicting the outcome.[17] These findings, much like the imaging and clinical studies evaluating WAD, tend to find small and inconsistent effects on the measured outcomes.[17] Other psychological factors, such as PTSD, have been found to have a moderate effect size and are deemed to be a significant risk factor for a poorer prognosis in a person with WAD.[103] It is unknown why these results vary so much, but it may be due to an incomplete understanding of the interaction of psychological factors and other bodily systems. To understand why some factors have a stronger influence than others, it may be necessary to look beyond the variables themselves.

Psychological factors are constructs that are not directly observable and are associated with signs and symptoms that can be physiological (i.e. weight loss) or psychological (i.e. feelings of sadness).[47] Certain factors may be responsible for some of the signs and symptoms associated with WAD, for example, the stress response system may play an important role in the mediation between psychological factors and chronic WAD symptoms.[64] In the stress response system, in particular the HPA axis, the release of hormones such as cortisol and catecholamines may influence the pain experience of the person.[40] The hormones can increase the sensitivity within the peripheral nervous system (PNS) as well as produce changes within the central nervous system (CNS),[36,40,65] potentially resulting in an atypical pain experience or the development of chronic WAD symptoms.

Psychological disorders occur on a spectrum,[47,104] some people demonstrate mild symptoms that barely rise to the level of the standard diagnostic criteria, while others are severely affected and have their lives significantly altered due to their condition. The primary psychological factor found to have the strongest relationship on the recovery trajectories of

chronic WAD symptoms is PTSD,[17,64] which is a long term psychological stressor that has the potential to alter the HPA axis.[40]

PTSD is a primary psychological disorder that can be present following a traumatic or threating event that results in avoidance of activities, recurrent and persistent thoughts and feelings related to the event, and hyperarousal of their senses.[47] The extent to which PTSD can alter the HPA axis relative to other stressors, such as depression or anxiety, or why PTSD is more strongly related to a poor prognosis with WAD is currently unknown. However, the argument can be made that PTSD has a stronger relationship because it is viewed as an immediate threat to the body and may increase the stress response system to a greater degree than depression or anxiety. With long term activation of this pathway and a consistently elevated stress response system, including sleep disturbances,[47] perpetual PNS and CNS activation can create increased sensitivity and symptoms.[17] For some people, this pathway may be enough of a stimulus to generate long-term WAD symptoms, whereas for others it may not.

The various pathways in which psychological factors can affect the WAD experience offers a deeper understanding of the BPS presentation in patients. These variables are not meant to be evaluated on their own as all symptoms originate from an MVC which can stimulate the peripheral and central nervous systems as well as psychological stressors. Thus, because both physical and psychological factors are related to the prognosis of those with WAD, the interactions between them must be understood in order to fully comprehend the condition and assist in improving outcomes.

2.4.4 Social Factors

For as long as researchers and clinicians have been studying and treating WAD, social factors have been thought to have an influence on its presentation,[46] however, no variables have

been found to have a strong relationship to symptoms.[17] Not dissimilar to the psychological variables, they have generally been evaluated in isolation and their effect sizes tend to be small and inconsistent.[17] Instead of evaluating the factors individually, it is possible that via mediation with other factors that their existence can be a factor in chronic WAD with a stronger association than has been previously reported.

For years, the notion of WAD as an injury has been questioned by authors of books and articles who present arguments against it as a legitimate pathology.[66,105] The concept of WAD being a falsified injury is advanced by the lack of findings on imaging studies, inconsistent signs and symptoms presented by individual patients, and even by the lack of a gold standard diagnostic test. Arguments that reinforce the idea that WAD is not a real injury tend to discount what the victim knows about their body; they hurt, and something is wrong. With cynicism regarding their injury, the person has to prove they are truly injured, and when their focus is shifted from learning how to cope and heal, to defending themselves regarding their injuries, very few will recover.[106]

Questioning the legitimacy of the injury is strongly encouraged by the "whiplash culture" theory.[66] The term whiplash culture was used to describe and understand why developed countries such as the US, Australia, and the UK report an almost epidemic of WAD, whereas other poorer countries such as Lithuania, and Greece have almost no reported incidence of WAD.[66] Beyond questioning the statistical and research design of the studies to explain the findings,[55] several theories attempt to make sense of the purported whiplash culture, even though it is unknown what the exact mechanism is for the differences between the countries.[66] There is the belief that it is as simple as the people of some countries do not recognize a whiplash event being a mechanism that can create chronic symptoms, and without this belief they do not have

cases of chronic WAD.[107] Others feel the drive for compensation and litigation, which is more prevalent in developed countries, can increase the rates of prolonged symptoms.[108] Finally, some see the system itself as an issue where the health care providers, insurance companies, or attorneys do not believe the patient is injured and thus patients spend all of their efforts proving they have been harmed rather than focusing on their own recovery.[106]

Regardless of the underlying reasons for the varying results between the countries, the studies that have evaluated these cross-cultural differences often come to similar conclusions: because the signs and symptoms of WAD are not identical across all cultures, then the premise of WAD as an independent injury is false and the disorder cannot truly exist.[66,109,110] With the differences in presentation, it is thought WAD is a result of the environment the person lives in, and it is not related to any specific biological injury or trauma sustained in the MVC. With this singular viewpoint of WAD, it is feared that countless people have had their symptoms marginalized and their injuries have been discounted which has resulted in increased suffering, pain, and prolonged disability.

A primary issue in these studies is the viewpoint that WAD is a homogenous condition that manifests similarly across all people rather than recognizing the heterogeneity and multiple variables that may influence the individual WAD experience. An argument can be made that these cross-cultural studies, instead of showing WAD as a single entity, demonstrate the heterogeneity in WAD and the influence of BPS factors. They accomplish this by showing the differences in various cultures and settings and how different social factors can affect outcomes.

Some of the previously described biological factors, age and BMI, can also be recognized as social factors that may influence WAD outcomes. As we age, there is often increased depression as well as issues with chronic pain which may impede a person's ability to exercise

and be active and this may impair the body's ability to heal.[99] As well, an elevated BMI may have a negative social stigma which can result in various psychosocial issues including but not limited to depression and anxiety,[111] both of which have been associated with chronic pain.

Another factor that is biological that also needs to be examined from the social perspective is the sex of the patients. Most reports on WAD that evaluate sex as a prognostic variable generally show that females are more likely than males to develop persistent symptoms, but this association usually ranges from weak to non-existent.[17] From a biological perspective, females are generally smaller than males and have less muscle mass, and the decreased muscle mass may be a predisposing factor in their higher rates of WAD than their male counterparts.[112] However, on the social side, car seats are generally designed to fit the average sized male driver, thereby the safety features of the seat may be inadequate to minimize the forces delivered to a female occupant.[112]

Therefore, it is not necessarily female sex, increased BMI, or age that increases the rates of WAD, instead, it is factors associated with these variables, biological or social, that may predispose someone to develop chronic symptoms. Because it is not uncommon for these factors to have multiple avenues in which they may influence outcomes, we need to ensure that we fully evaluate them for their individual and combined contributions to persistent symptoms.

Other social constructs that have been researched are the education and socioeconomic level of the injured person.[113] However these factors, much like the previous ones, have inconsistent findings with small to non-significant effect sizes.[17] With variables like these, there are potentially many mediating factors that can be related to a poorer prognosis. Social factors such as missing medical appointments due to an inability to leave work or attain childcare may influence an outcome. These factors, especially when combined with the influence on pain by the

neurophysiological and psychological variables previously described, may demonstrate stronger effect sizes than have been previously reported.

2.4.5 Sleep and Chronic Pain

Sleep is an activity that all people partake in, we spend approximately one-third of our life in a sleep state, but its purpose is not fully understood.[51,114] Sleep is believed to be a key to maintaining homeostasis across multiple bodily systems[115] including memory consolidation,[114] brain development,[51] cellular repair,[51] and facilitating immune functions.[116] It is likely that sleep has some capacity to affect all of these functions, and a lack of sleep may potentially facilitate adverse consequences.

Lack of sleep, in particular insomnia defined as 30 minutes or more of sleep latency 3 times or more per week for 3 or more months, has been found to have a relationship with chronic pain.[117] In particular, it was reported that increased nociceptive hypersensitivity and the low grade inflammatory reactions were a consequence of decreased sleep and may be a contributor to the development of chronic pain.[116] Thus, the amount of sleep interference plays a role in a person's chronic pain with a small and significant effect.[118]

Historically, the relationship between chronic pain and lack of sleep has been viewed as bidirectional as it was believed that each one equally influenced the other.[115,117] However, a review by Finan reported that the relationship may be more strongly associated with sleep affecting pain than pain influencing sleep.[115] Although the literature is quite extensive on the relationship between chronic pain and sleep disturbances, the evidence between chronic WAD and sleep is not fully developed. Sleep disturbances have been found to be prognostic of poor outcomes at 4 and 12 weeks post MVC, but the amount of explained variance is small ranging from 2.9% to 7.8%.[119] Another study found that those with WAD had significantly greater

disturbances in sleep when compared to those with insidious onset mechanical neck pain.[120] As well, in those with WAD, a statistically significant correlation between pain and sleep quality (r = 0.693, no CI provided) and sleep duration (r = 0.433, no CI provided) was discovered; there was no significant correlations with those variables for people with insidious onset neck pain.[120] Thus, sleep may be recognized as a small but significantly important factor in the patient with chronic WAD symptoms.

2.4.6 Tissue Trauma and Imaging

Damage to tissues in the neck following a whiplash incident has been observed in cadavers,[8] animal models,[26] and computer simulations,[70] and the results from these studies are consistent in their findings. However, imaging studies on patients with WAD have been unable to corroborate consistent results.[5,24,25] When imaging studies do find abnormalities, they are either not significantly different from a non-injured control group, or the findings are so rare they cannot explain the large numbers of people with WAD.[24] Yet, it may be argued the findings are rare due to the inconsistent protocols used across a number of imaging studies.[7]

Reading radiological images can be a difficult endeavor given the various shades of gray and illuminated voxels making up the images, especially when combined with the multiple sequences available for viewing the anatomy.[10] These shades of gray represent the anatomy of the patient and based on the reading of the radiologist can represent a normal or pathological finding. Reading images does not always provide consistent results, for example, a study by Anderson evaluated 4 different readers of 10 separate pathologies on 200 subjects with WAD and reported a 0.328 sensitivity and 0.728 specificity for all combined pathologies.[24] According to these result and the reported sensitivity, imaging is not efficient at ruling people out for the development of chronic WAD. It is this group of patients that may not be appropriately managed

following the MVC as they may be requiring care, but with a negative test result, they are not receiving it.

It is not fully understood why there was a low sensitivity; one theory is that it may be a result of the limited scoring flexibility within the protocol. For example, in the Anderson study,[24] they attempted to quantify various pathologies with a dichotomous result, a yes or no for having the pathology. Disc degeneration was defined as a loss of signal intensity on T2 weighted spin echo images, and high grade foraminal stenosis had no specific criteria listed.[24] However, given there are varying degrees of disc degeneration and the concept of high grade foraminal stenosis was not clearly defined, it is possible this lack of criteria led to inconsistencies and low sensitivity. Therefore, a scale that labels pathologies a dichotomous variable may not be useful in the interpretation of images as the results can be inconsistent and perpetuate the notion that imaging is of little use in assessing damage in the neck following an MVC.

To avoid these ambiguities, it can be helpful to have set criteria that allows the radiologist the opportunity and flexibility to read, interpret, and report based on what they are seeing. For example, a study by Pfirrmann evaluated the lumbar spine in which a particular scale was established for disc degeneration ranging from 1-5 with specific criteria for each level.[121] The authors reported the intrarater reliability to be excellent (Kappa, 0.84 – 0.90) and interrater reliability was substantial to excellent (Kappa, 0.69 – 0.81).[121] Also using this scoring criteria during an MRI of the lumbar spine, researchers reported those with a grade 3 or above had a higher likelihood of developing back pain when compared to those who had a grade 2 or below.[122] With these results, it is possible that not all disc degeneration is the same, and the need to reliably and consistently grade the discs (or any anatomy) may be an important factor in determining pathology.

Beyond the scoring criteria and subjectivity of the readings, other theories exist about why imaging studies are unable to find consistent pathology following a whiplash trauma. The first idea has to do with the protocols followed regarding the parameters for the machine. A second theory is related to the role of technology and whether the machines are sensitive enough to detect the subtle damage that can occur with an MVC.

To establish internal validity, a consistent protocol must be created from two primary components: the parameters of the machine and the readings by the practitioner. For the machine, there are numerous variables that must be accounted for when trying to determine the most optimal set up for acquiring images. Each variable has positive and negative attributes in its ability to improve the consistency and resolution of the images,[10] but they must be used repeatedly if co-registration of the images is going to be possible.[123]

After the protocol for the machine is established, procedures for those reading the images need to be consistent. The importance of identifying methods that demonstrate strong interrater reliability cannot be understated; if the results are not consistent from clinician to clinician then the test is of limited use. Therefore, having a standardized protocol or a checklist to assist with reading and interpreting images could be of great value.

It is not known if imaging technology is sensitive enough to detect injuries that can occur from an MVC. However, the results from this dissertation and the protocols used to obtain the images will hopefully assist in the standardization of future studies to allow for comparisons between various groups. With improved standardization, statistical methods such as meta-analyses can be used to answer the questions regarding anatomy and its relationship with prolonged symptoms.

2.5 Relevance

2.5.1 Diagnosis vs. Prognosis

The notion of creating a diagnosis in cases of a whiplash injury is something that needs to be reconsidered based upon the limited objective diagnostic markers available. Without a gold standard to make the diagnosis of WAD, and with little to no clinical benefit regarding a change to the patient's outcome based on this diagnosis, the idea of evaluating their prognosis based on several variables may be of greater clinical value.

Recognizing the three distinct pathways in which a patient can progress after they are injured, and the prognosis associated with each one, has the potential to guide and improve clinical care. The three pathways are: 1) no long term to symptoms 2) mild symptoms 3) moderate to severe symptoms.[124] Each one of these pathways can be influenced by any or all the previously discussed BPS factors. By having a greater understanding of all variables strongly correlated with the moderate to severe pathways, clinicians and patients can have a clearer picture of the patient's future.

As determining the proper prognosis is of great interest, a novel concept referred to as triangulation has been proposed to help understand where the patient is located in relation to their recovery pathway.[125] Using the process of triangulation to determine a prognosis is analogous to using multiple cellular towers to determine the location of a cell phone (where each WAD variable is equivalent to a different tower). If only one variable is used to determine the patient's prognosis, then all outcomes related to that factor are possible. However, if two variables are used to determine the prognosis of the person, then wherever there is overlap between the two factors is where the person is located along the recovery trajectory. The second variable greatly reduces the number of potential outcomes available when only a single variable

is used. Greater precision can be accomplished when a third variable is used, as the person's location is further developed based on the overlap of all three variables.

2.5.2 Vulnerabilities and Resilience

In addition to the BPS factors that can influence the prognosis of WAD, the timeline of WAD must be examined to fully comprehend all the possible influences on symptoms. In general, the progression of WAD is viewed on a continuum where the MVC provides a starting point for the signs and symptoms and hopefully a resolution. However, there is increasing evidence that factors present prior to the MVC can affect prognosis and the recovery trajectory.[30] These pre-MVC factors are referred to as vulnerabilities,[30] and although they are unrelated to the MVC in the sense that they were not caused by the collision, their presence may have a significant effect on the outcome.

The vulnerabilities are part of the diathesis stress model, where the term diathesis refers to the increased tendency or susceptibility to a particular medical condition.[30] The diathesis can be based on any or all of the potential vulnerabilities and when combined with a particular stress, an MVC for example, can lead to a maladaptive pain response. Given each person has their own set of vulnerabilities and the pain experience is individual, it is understandable why researchers have been unable to create a single standardized picture of those afflicted with WAD.

Vulnerabilities can come in various forms including but not limited to psychological, physical, and genetic.[30] Any of these variables, and most likely it is a combination of them, can lead to a maladaptive response following an MVC and can influence the recovery pathway leading to chronic moderate to severe WAD. Therefore, being aware of these vulnerabilities may be of value in helping to determine the person's prognosis.

For those diagnosed with WAD, several variables have been found to have moderate to strong associations with the transition to chronicity (Table 2.2). Some of these factors may be attributable to the MVC, such as the initial pain rating or their NDI score (in the sense that had they not been in the MVC then they may not have any neck pain or an elevated NDI score). On the other hand, factors such as quantity of MFI may be present prior to the MVC based on previous experiences or even genetics. Because of these pre-MVC vulnerabilities, some people may be at a greater likelihood of developing long-term symptoms and if they can be discovered early in the process then perhaps their treatments can be modified to enhance recovery.

When evaluating the potential for treatment, some factors such as age, sex, or socioeconomic status are non-modifiable, whereas other factors such as a high initial pain score or elevated initial NDI score cannot be easily modified if they can be changed at all. Although the pain rating and NDI score are strongly associated with a recovery trajectory, without a clinically feasible method to alter these variables, this knowledge does little to changer a person's treatment or their prognosis. However, other factors are modifiable and if changed could lead to improved outcomes, for example, MFI has the potential to be modulated with specific exercises,[96] and PTSD can be managed with psychological or educational interventions.[126] It is unknown if treatments focused at these variables for a specific subgroup of people with WAD can decrease pain or the rates of disability.

With the interconnectedness of the BPS model and the heterogeneity of each person, it must be recognized that treatments will not be as simple as addressing individual issues or symptoms. Instead, all BPS factors, as well as pre-existing vulnerabilities may need to be considered, and their relationships with each other will be evaluated to determine and help manage their recovery trajectory.

Table 2.2 Prognostic Variables and the Strength of Relationship to Chronic WAD[17,20]

Prognostic Variable	Strength of Association
High initial pain intensity; greater than 5.5/10	Strong
High disability score; NDI greater than 14.5/50	Strong
Presence of Post-Traumatic Stress Disorder	Moderate
Elevated quantity of MFI	Moderate

Key: WAD – Whiplash Associated Disorder; NDI – Neck Disability Index; MFI – Muscle Fatty Infiltrate

2.6 Underdeveloped Factors Related to Chronic WAD

Even with the current knowledge regarding prognostic variables and recovery trajectories related to WAD, a significant amount of information is still unknown. The inability to generate meaningful clinical information may be due to prognostic variables being studied in isolation with the results focused on how individual factors influence outcomes at a specific point in time. With all this specificity, it is not understood how these factors interact with one another or how these relationships can change over time. For example, the stress response system can vary over time as some variables can take weeks to months to fully manifest which means the effect of these interactions will fluctuate. With this changing influence, the relationships must be understood throughout the entire timeline of the WAD experience if we are able to provide a meaningful understanding of the condition to the patients.

The influence of psychosocial factors and how they can interact with biological changes and other known prognostic variables is not fully developed in the literature. In addition to these BPS variables, there are pre-existing vulnerabilities that may predispose someone to developing WAD. Therefore, evaluating and understanding how these vulnerabilities can influence outcomes, and how they can be measured and put into a model to assist clinicians and patients to understand the prognosis may be of great value. By recognizing the heterogeneity of all patients

and the various factors that can contribute to their symptoms, we may begin to have truly individualized treatments that can focus on minimizing the transition to chronicity.

The biological underpinning of WAD has been addressed in numerous studies, but conclusive results continue to evade patients, clinicians, and researchers. With inconsistent protocols between studies, and an understanding that a single lesion for all WAD symptoms is unlikely to be discovered, there is the increasing need to determine the role of structural changes in the tissue and how it can affect symptoms. With improved technologies and structured protocols, changes in the tissue such as the deposits of MFI may be more readily recognized as playing a part in the manifestation of WAD. By recognizing multiple variables that may influence WAD, an aggregate score of these factors may be of greater importance than trying to isolate a single lesion to explain the symptoms.

2.7 Contributions to the Literature

The crux of this dissertation was to evaluate MFI and its relationship to various BPS factors regarding the incidence and recovery of WAD at baseline and across a 12-month timeline. The detection and grading of MFI on MRI was evaluated and related to psychosocial factors associated with outcomes in a heterogenous population enrolled in a separate NIH-funded parent study (RO11HD097076). From this study, the theoretical construct of MFI that occurs due to an MVC and its influence on symptoms was examined.

Initially, the study focused on understanding how MFI may be related to the transition to chronicity based on its presence at baseline. The understanding was evaluated by recognizing the association between MFI and NDI. Once this relationship was established, an evaluation of how the relationship between these two variables changed over a 12-month period was carried out.

Evaluating MFI as a single predictor of explaining WAD signs and symptoms will more than likely will not yield clinically meaningful results. As was previously stated, known BPS factors should be considered to understand how and why a person with WAD can transition to the moderate to severe pathway. The variables that may influence the signs and symptoms associated with WAD are individual and can change over time, and the changes in these variables may explain the variety of symptoms that people with WAD often describe. Therefore, a second phase of the dissertation will be to explore how the relationship between MFI and NDI over time is altered when considering known influential BPS factors.

The results from this dissertation will benefit in developing a more complete understanding of MFI and the biological sequela of an acute injury through the development of chronic WAD and the relationship with known prognostic variables. To accomplish this, multiple BPS domains and their interactions with other areas need to be evaluated to determine which areas explain the greatest variance of WAD symptoms. When their interactions are understood, future research and treatments can be focused at modifying these variables to help reduce the transition to chronicity.

Chapter 3: Methodology

3.1 Introduction

The purpose of this chapter is to describe the methods used for data collection and detail the statistical analysis that used to generate results. Recruitment procedures, inclusion and exclusion criteria, demographics, and variables collected for study participants are included. Protocols used to acquire imaging data, including user prescribed parameters to ensure quality, as well as the methods used by researchers to determine the quantity of muscle fatty infiltrate (MFI) are described. A description of the self-report outcome measures collected at each time point as well as psychometric properties, including reliability and validity for each outcome measure are provided. Statistical methods to discover potential relationships between known prognostic variables, the quantity of MFI, and the participants recovery pathways are described, and these methods were used to help determine results for each specific aim of the study.

3.1.1 Specific Aim #1

The first aim of this study is to determine the combined relationship between baseline MFI values and prognostic variables believed to influence the outcome of participants diagnosed with whiplash associated disorder (WAD) on neck disability index (NDI) scores.

3.1.2 Specific Aim #2

The second aim of this study is to determine if there are changes in magnitude in the strength of these relationships over a 12-month period.

3.2 Research Design and Outcomes

3.2.1 Research Design

Data for this study was collected as part of a National Institute of Health (NIH) funded research project clinicaltrials.gov Identifier: *NCT02157038*. From this study, a secondary analysis of the data was performed to discover the results associated with the previously described aims.

The NIH project was a longitudinal observational study in which participants were recruited immediately (within one week) following a motor vehicle collision (MVC) and followed over a 12-month period. During those 12 months, magnetic resonance imaging (MRI) of their cervical spine was obtained at 1-week, 2-weeks, 3-months, and 12-months post-MVC. In addition to imaging, enrolled and consented participants were invited to complete a variety of self-report measures at each time point evaluating their functional impairments, signs of post-traumatic stress, anxiety and depression, distress associated with the MVC, and pain intensity. Although no treatments were offered to the participants, they were not restricted from seeking and receiving any treatment for any MVC related complaints.

3.2.2 Image Acquisition

Images of the cervical spine were acquired using a 3-Tesla (T) Siemens MRI scanner. For the scan, a 3D-multi-echo, Dixon fat/water separation gradient echo approach was used to quantify the total percentage of MFI.[127] The 3D-multi-echo method has been found to be as reliable as a T1 weighted image in determining the amount of MFI present in the musculature[128]. However, the 3D method acquires the images at a significantly faster rate thereby decreasing the possibility of image artefacts from the participant moving and affecting image quality.[127]

During each of the four MRI examinations, images were obtained from the most superior portion of C3 to the most inferior portion of C7. To maximize co-registration, a standard 12-head channel coil and neck specific receiver coil were used to obtain images. Previously established parameters for obtaining the images related to maximizing the distinction between fat and water included using a T2-weighted sagittal turbo spin echo sequence with a TR of 23.81 ms, 8 echo times with a spacing of 1.78 ms starting at 1.36 ms, as well as a slice thickness of 3mm for each image with 22% overlap to prevent aliasing.[127] Finally, a rectangular field of view of 75% with a 1.4 mm in-plane resolution was used to view the area. Regions of interest of MFI were drawn with Analyze Software (v. 11) at the most superior portion of each vertebrae from C3-C7 for bilateral muscle groups: multifidus, semispinalis cervices, semispinalis capitis, and splenius capitis. Bilateral muscles were evaluated and used to determine the total amount of MFI.

Using the Dixon method, the images can show water or fat only, as well as water and fat combined. This ability to isolate each one makes it possible to determine the amount of MFI present. The following formula was used to calculate total amount of fat as a percentage:

$$\text{Percentage of Fat} = 100 * \text{fat} / (\text{fat} + \text{water})$$

The amount of MFI was measured and recorded by a single researcher blinded to the outcome of the participants. The information from the researcher's measurements, data from the self-report measures, and the participants demographic information were transferred to a Microsoft Excel (2016) spreadsheet for analysis.

3.2.3 Sample Size Estimation

The estimated sample size for this study was determined to be 100 participants with the assumption based on previous research that 27 will develop chronic moderate to severe

symptoms, and 63 will fall into the recovered/mild category.[4] The one-hundred participants allows for a 10% dropout rate yet still provides sufficient statistical power for data analysis based on an 80% power to detect a change in MFI of 0.66 standard deviations between the groups and a Type I error significance level of 0.05.

3.2.4 Recovery Pathways

Recovery pathways, or predicted recoveries, are classification schemas that attempt to determine homogenous subgroups of people who are likely to report similar outcomes over a period of time.[4] The pathways are determined by group based analytical techniques in which participants are followed and the outcome of interest is measured over time.[4] With similar outcomes, participants are grouped together based on their recovery pathway and common variables are examined to determine the influence of these factors on the trajectory.[4]

In those with WAD, three distinct recovery pathways based on NDI scores have been reported: 1) recovered $\leq$ 10%; 2) mild symptoms 10 to 28%; 3) moderate to severe symptoms greater than 30%.[4,124,129] Several studies have used different outcome measures to evaluate the presence of these pathways,[4,129] and they collectively report three distinct groups. The NDI is the most frequently cited outcome measure associated with these pathways. (Table 3.1)

While these pathways may provide prognostic value associated with long-term recovery rates, the greatest amount of recovery occurs during the first three months following the MVC with little to no improvement after that timeframe.[4,124] The lack of improvement after three months may have clinical implications because it is not uncommon for a person with WAD to delay initiating a formalized rehabilitation program. Therefore, if there is limited improvement after three months then the initial delay may limit their recovery.[129] For this reason, having a clinical tool that allows practitioners to predict (or at least increase their confidence) which

trajectory a person may follow could be beneficial to focus resources and treatments towards subgroups that may have the greatest potential to transition to chronicity.

Table 3.1 Recovery Pathways Based on NDI Scores[4,124]

Group	NDI score
Recovered	**0-8%**
Mild symptoms	**10-28%**
Moderate to Severe symptoms	**>30%**

NDI: Neck disability index; scored as a percentage.

3.2.5 Institutional Review Board

As this is a secondary analysis of data collected through an NIH study, *Neuromuscular Mechanisms Underlying Chronic Whiplash*, the institutional review board (IRB) of record for that study is Northwestern University (NU) in Chicago, Illinois (IRB# STU00040759). However, the current PhD dissertation study was run through Nova Southeastern University (NSU), so prior to execution of the primary actions of this study including the use of data for computations and analysis IRB approval was evaluated. As this was a secondary analysis, it was deemed that this study was exempt from the IRB approval process. (Appendix A) In addition, as the data was collected at NU, a data sharing agreement between NSU and NU was obtained. (Appendix B) It is also recognized that NSU will oversee the actions of this PhD dissertation but no other past, ongoing, or future, study activities or outputs from the NIH parent study will be through this university.

3.3 Participants

3.3.1 Inclusion and Exclusion criteria

The primary factors for inclusion in the study was an MVC within the past week which resulted in a diagnosis of WAD. For the diagnosis of WAD, participants had to meet the criteria for the Quebec Task Force (QTF) grades 2 (neck pain, decreased range of motion, and tenderness) or grade 3 (same as grade 2 in addition to neurological changes in reflexes, sensation, or strength). Potential participants had to be between the ages of 18-55 and both males and females were eligible.

A participant was excluded from the study if they had a fracture of the spine related to the MVC, a previous MVC, history of spinal surgeries, or prior diagnosis of cervical or lumbar radiculopathy. A history of neurological disorders (e.g. multiple sclerosis, Alzheimer's, stroke, or myelopathy) inflammatory diseases (e.g. hepatitis, systemic lupus erythematosus, rheumatoid or osteoarthritis, ankylosing spondylitis, Chron's disease, fibromyalgia) or metabolic diseases (e.g. diabetes mellitus, hyper or hypo thyroid) eliminated potential participants from the study. Inability to tolerate (claustrophobia) or be medically cleared (metal implants, pregnancy, or a pacemaker) for the MRI procedure also rejected potential participants from the study.

3.3.2 Recruitment

The participants were recruited from the Chicago area, primarily through the emergency department (ED) at Northwestern University (Northwestern Memorial Hospital) as well other EDs in the surrounding region. Potential participants were contacted either in the ED or by a phone call the following day by a member of the research team to determine their level of interest, explain the general procedure of the study, and expectations for participation. To recruit participants that did not seek medical attention, a flyer was posted on social media that listed the

inclusion and exclusion criteria for the study. Interested parties could contact the research team via email where they could state their desire to participate. Upon follow-up communication, or in person in the ED, a brief screening examination reviewing the inclusion and exclusion criteria was performed with each potential participant to determine their suitability for the study.

3.3.3 Demographics

Patient demographics such as age, sex, height, weight, race, socioeconomic level, education level, and litigation status were collected via a standardized intake form (Appendix C). A basic health history form (Appendix C) was also collected to determine any potential medical complications, headaches, or injuries that existed prior to the MVC.

3.3.4 Informed Consent

All participants who were candidates and elected to participate in the study were required to sign an informed consent form (Appendix D) which outlined their rights as a research participant, potential risks associated with their participation, and the expectations placed upon them. They were informed that their participation was voluntary, and they could withdraw at any time.

3.3.5 Privacy Protection

The collection and storage of all participants information was treated and protected with a standard protocol as established by the Health Insurance Portability and Accountability Act (HIPAA). Once collected, all original forms were scanned into a computer for storage and all hard copies were stored in a secured cabinet. Prior to inputting the data into the Excel spreadsheet, the participants name and identifiable information was deidentified by way of creating a participant ID number. All digital information was secured on a password protected computer.

3.4 Self-Report Measures

3.4.1 Neck Disability Index

The NDI is the most widely used self-report outcome measure for determining limitations due to cervical spine pain.[48] The NDI is a 10-item survey with seven items assessing activities of daily living (ADL), one for neck pain, one for headaches, and one related to concentration; the items are scored on a 0-5 scale with 0 meaning no disability and 5 representing full disability.[130,131] (Appendix E) There are two ways to score the NDI: 1) The raw score provided by the patient is out of 50 points if all 10 questions are answered or fewer points if they do not answer all of them; 2) The participants total score is divided by the total points possible to generate a percentage of disability.[130] The percentage of disability method works well when questions are not answered on the form,[48] for example, someone may not drive and therefore cannot provide a response to the question that asks about symptoms while driving. However, if three or more items are not answered then the tool is considered invalid.[132] For this study, the percentage of disability will be used for analysis.

There is debate as to the number of dimensions that are evaluated by the NDI. The intent of the NDI was to create a unidimensional scale to measure the functional status/disability of a person with neck pain related to their ADLs.[130] However, it has been proposed that the NDI may evaluate two separate dimensions: 1) functional disability and 2) pain and interference with cognitive functioning.[133] Although the number of dimensions that are evaluated is not agreed upon, the NDI is considered an adequately valid, reliable, and responsive outcome measure for people with neck pain.[48]

The NDI has been used across multiple subgroups of people with neck pain including those with WAD.[48,133,134] Specifically for those with subacute WAD, the reliability of the NDI

has reported an internal consistency with a Cronbach's alpha 0.87,[133] which is considered acceptable.[135] In a systematic review, the authors reported the test-retest reliability correlation coefficient r ranged, based on acuity and time between measurements, from 0.73 to 0.93, and specifically for those with chronic neck pain r = 0.89 to 0.99.[48]

Instead of calculating a standard correlation coefficient, Cleland performed an intraclass correlation coefficient ($ICC_{2,1}$) which evaluates not only the degree of correspondence but also the agreement among the ratings.[136] With the ICC, the authors reported results 0.50 for people with mechanical neck pain (95% confidence intervals (CI), 0.25 to 0.67) and for cervical radiculopathy 0.68 (95% CI, 0.30 to 0.90).[131,137] Although the results reported by Cleland are not as robust as previous authors, the test-retest reliability of the NDI is still considered fair to moderate and is deemed acceptable to use in both clinical and research settings.[48]

The validity of the NDI has been studied extensively and in general it is found to have good to excellent results.[48] The convergent validity between the NDI and other outcome measures has been evaluated and a correlation coefficient of 0.70 and greater has been reported.[48] These results infer that the NDI is acceptable to use for evaluating both pain and functional disability for someone with neck pain.

The responsiveness, both the minimum detectable change (MDC) and minimum clinically important difference (MCID) of the NDI has been reported with a wide range of results. The MDC is the amount of change that must occur to ensure that it has exceeded the standard error of measurement (SEM);[136] the lowest reported MDC for the NDI was 2/50 points,[138] while the highest was greater than 10 points out of 50.[131,139] The wide variation in scores related to the MDC may be due in part to the participants that were used or the

determination of what constituted a change in a person's status.[132,139] Even with this variability, the use of 5 points out of 50 for the MDC is generally deemed acceptable.[132]

A receiver operator characteristic (ROC) curve has been used to determine the MCID for the NDI. The ROC curve is a graphical representation of a balance between sensitivity and specificity that helps to determine a cut point based on the relative importance desired by the researcher.[139] From the ROC curve the area under the curve (AUC) can be calculated. The AUC is a measure of responsiveness represented by a single number as a probability of correctly predicting an outcome; an AUC of 1.0 is a perfect prediction and 0.50 is equal to simply guessing.[136]

For the NDI using a ROC curve, authors have reported a MCID of 7.0 out of 50 points with sensitivity of 0.52 (95% CI, 0.32 to 0.71) and specificity of 0.59 (95% CI, 0.36 to 0.78) and an area under the curve (AUC) of 0.57 (95% CI, 0.38 to 0.75).[131] A second report of the MCID using a ROC was 3.5 points out of 50 with sensitivity of 0.9, specificity of 0.7; neither the CI or AUC were reported.[139] Although there is a spread of points for the MCID between these two studies, the use of 5 points as the MCID in studies tends to be accepted.[132] However, some subgroups may be more resistant to improvement and may therefore require greater changes in scores to generate a meaningful MCID.[132]

One other area in which the NDI has limitations is in the floor and ceiling effects. The NDI can have difficulties in its responsiveness when a person is near the extremes of the scale.[48] For example, if a person has low disability, 3/50, it may be difficult for the NDI to capture improvements made by the person. In addition, the validity, reliability, and responsiveness have not been assessed at various scores (low, medium, and high) as the scale may act differently throughout the spectrum of possible responses.[140]

Although the NDI is deemed appropriate to use as a measure of functional loss in those with neck pain, it may not capture all the domains that are important to those with WAD as they often present with emotional or social issues that may complicate their presentation.[134] Therefore, using multiple measures to evaluate a person's function and emotional health may be of benefit in accurately describing their complete presentation.

3.4.2 Post-Traumatic Stress Disorder

PTSD is a mental health disorder that manifests with psychological distress after a person is involved in, or witnesses, a traumatic event, including an MVC.[47] From the traumatic experience, three primary domains of their symptoms will arise: 1) The person may develop recurring thoughts or feelings about the event (intrusion); 2) They may avoid situations that they fear may trigger their memory of the event (avoidance); 3) They tend to be always on alert or looking out for danger (hyperarousal).[47] With these symptoms, many people find it difficult, if not impossible, to manage their everyday activities, which can lead to further mental and physical health issues. For this study, symptoms associated with PTSD were measured using the Post Traumatic Diagnostic Scale (PDS).

3.4.2.1 Posttraumatic Stress Diagnostic Scale

The Post-traumatic stress diagnostic scale (PDS) is a self-report questionnaire that is designed to measure the severity of PTSD symptoms related to a specific incident or trauma and assist in generating a diagnosis of PTSD.[50] For data collection related to this study, the MVC was considered the traumatic event of record. The PDS consists of 49 items with 17 specifically about the symptoms associated with PTSD and the other 32 about the traumatic event and the interference of the symptoms on activities.[50] The PDS has been found to demonstrate adequate validity and reliability for the assessment of the severity of PTSD symptoms.[50]

The internal consistency of the entire scale has been estimated at $\alpha = 0.92$.[50] The test-retest reliability has been calculated for both the generation of a diagnosis of PTSD and the symptom severity. For the PTSD diagnosis and symptom severity, between 2 and 3 weeks, kappa values for the PDS have been reported as 0.74 and 0.83 respectively.[50]

For the diagnosis of PTSD, the results of the PDS have been compared to the structured clinical interview which is the most commonly used criterion on which to validate a PTSD measure.[50] For convergent validity of the PDS and the interview, a kappa value of 0.65 with sensitivity of 0.89 and specificity of 0.75 have been reported which indicate satisfactory agreement between the two measures.[50] The symptoms of PTSD as measured with the PDS correlated with the Impact of Events Scale-Revised subscales of intrusion and avoidance with $r = 0.77$ and 0.69 respectively which also demonstrates adequate agreement between the scales .[50]

The full PDS score was not a part of this study but will be used within the parent NIH study. In this dissertation, only the score from the PDS related to sleep disturbance was used as an isolated variable.

3.4.3 Numeric Pain Rating Scale

The Numeric Pain Rating Scale (NPRS) is a commonly used tool to determine the level of pain a person is experiencing.[141] The NPRS is an 11-point scale that ranges from 0 to 10 with scores on each whole integer.[141] The anchors for the scores are no pain at 0 and the worst pain imaginable for the score of 10.

The psychometrics of the NPRS has been reported across various subgroups and in general it is found to be a valid, reliable, and responsive tool.[137,142] The validity of the NPRS has been established by correlating it with the Visual Analog Pain Scale, of which $r = 0.94$ (95% CI,

0.93 to 0.95) was reported.[142] In a group with mechanical neck pain, the 3-day test-retest reliability as measured with an $ICC_{2,1}$ was reported as 0.76 (95% CI, 0.52 to 0.87).[137] For determining the responsiveness of the NPRS, the standard error was 0.91, with MDC of 2.1, and 1.3 for the MCID.[137] When the NPRS was correlated with the self-report global rating of change, the 1.3 MCID demonstrated a sensitivity of 0.88 and specificity of 0.71.[137]

3.5 Data Analysis

3.5.1 Descriptive Statistics and Data Screening

All data in this study were analyzed using SPSS version 24.0. Prior to analysis, the data were screened to assess for normal distribution. Participants were placed into one of two groups based on their NDI score at 12-months: 1) Recovered less than 10%, 2) Persistent symptoms greater than or equal to 10%. Group membership was based upon findings from previous studies determination of recovery trajectories.[103]

For all variables the skewness and kurtosis were reported, and for those with multiple time points it was reported at each one. A value of zero will be recognized as normally distributed for both skewness and kurtosis. However, there is no consensus on how much variability is acceptable before data is determined to skewed or kurtotic.[143] Within the output for each variable, any number above 1 or below -1 will be deemed skewed or kurtotic.[143]

To assess normality for each variable using a numerical method the Shaprio-Wilks test was employed. A significance value below 0.05 was considered a departure from normal. Graphical representations of normality were assessed using histograms, Q-Q plots, and box plots to determine the spread of the data and determine if any outliers are present. If any outliers were discovered, their data was handled as follows: 1) the original data was reviewed to ensure it was

inputted properly 2) if there were no errors their chart was reviewed to evaluate if there are any reasons (i.e. medical conditions) that may have influenced the outlier status and a determination was be made about retaining or discarding their data.

Baseline characteristics between the groups was established for each variable. At the interval/ratio level an independent t-test was used, a Mann-Whitney U test was used for ordinal level data, and for nominal level data, a chi-square (χ^2) test was calculated to determine if differences in proportions exist between the groups. An alpha level of 0.05 was used to determine significance for all group differences.

For missing data, the individual cases were handled in one of two ways. If the data that was absent was related to a previous time point, similar to an intention to treat analysis, the last measured number was carried forward to the missing variable. If the data missing was related to a non-repeated measured variable, for example age, the mean for the final group membership in which the participant is a part of was used in its absence.

Finally, multicollinearity of the variables was assessed using a correlation matrix. As there were interval/ratio level data as well as ordinal level a Spearman Rho correlation was used to evaluate the variables. Any r value at or above 0.75 for two variables was considered to demonstrate multicollinearity.[143] If any two variables demonstrated an r value at or above the designated level, based on careful consideration and reasoning, one of the two was removed from the analysis.

3.5.2 Aim #1

A logistic regression analysis was used to determine the relationship between the dependent variable (DV), group membership, and the independent variables (IVs) previously

listed for the baseline data. Interaction effects between the variables that were found to be different between the groups at baseline were also included as IVs. Prior to entering the variables in the model, a decision was made on how to handle ordinal level data. Logistic regression is adept at managing categorical and continuous level data, bur ordinal level data must be classified as either categorical or continuous in order to be properly analyzed.[144] Therefore, each IV that was recorded at the ordinal level was evaluated via statistical methods as well as using logical reasoning to determine which path, categorical or continuous, the variable would be categorized.

Once all variables were established, a backwards conditional method was used for determining which IV's to include in the final model, a p-value of ≤ 0.10 was used as the alpha level for inclusion into the model. From the final model, the Nagelkerke R^2 was reported, as well, for each variable the unstandardized beta weights, odds ratios (OR), and 95% CIs for the beta and OR were stated. In addition, sensitivity and specific were calculated along with ROC and AUC to assist in determining group discrimination.

3.5.3 Aim #2

Using only the variables that were found to be significant at baseline, separate analyses were run for each subsequent time point described in section 3.2.1: 2 weeks, 3 months, and 12 months. To evaluate the influence of the IV's at each time point during the 12-months following the MVC, the OR and its CI were compared to one another to determine if the strength changed in magnitude or direction. A graphical representation of time and the OR and CI were evaluated and compared at all time points to determine any potential changes for the variables.

3.6 Summary

This chapter establishes the methodology for this project including the research design, acquisition of data and participants, protocols followed, self-report measures used, data analysis carried out, and their potential limitations. Every attempt has been made to base all decisions on the most current evidence and to provide the most complete and thorough results.

Chapter 4: Results

4.1 Introduction

The purpose of this chapter is to report the results of the statistical analysis from the study –

Neck Muscle Composition in Persistent Whiplash Associated Disorder:

Relationship with Neck-Related Disability

The chapter describes the data cleaning and screening used to assess for normality; discusses the results for evaluating variables to include into the logistic regression model; reports on how the variables and model changed over time; and examines if there were any interaction effects between the variables. The statistical analysis was performed using SPSS version 24 (IBM Corp., Chicago, IL.)

The questions guiding this analysis were 1) What factors have a statistically significant influence on predicting neck-related disability in persistent whiplash disorder? 2) What is the magnitude of the influence for each variable? 3) How do these factors interact at baseline and potentially change over time?

Initially, the variables at baseline were assessed to determine which ones were significantly different between the groups with persistent symptoms and those who later reported they were recovered based on their 12-month scores from the Neck Disability Index (NDI). For evaluation of the data, continuous-level variables were examined using independent *t*-tests, Mann-Whitney U tests for ordinal level, and Chi-square tests for nominal level. Following these tests, a logistic regression model was built by assessing all variables and determining those that were statistically significant predictors of persistent symptoms; the individual strength and direction of each variable was also reported. In addition, the overall model prediction was stated

by way of the Nagelkerke R^2, sensitivity and specificity, a receiver operator characteristic (ROC) curve and the area under the curve (AUC)

The second group of questions evaluated the logistic regression model's prediction, as well as the variables within the model for their influence over pre-determined time points. The variables were assessed based on their unstandardized beta value, odds ratios (OR), and 95% confidence interval (CI). The OR and 95% CI were assessed to determine if a significant change occurred at any of the time intervals. Finally, interaction effects of the variables were assessed to determine if any group by time interactions occurred that may not have been visible viewing only the simple main effects.

4.2 Data Screening

Prior to running the analyses, the data were cleaned for missing records, screened for normal distribution, and assed for outliers. Currently, there is no single method used to determine normality, therefore, in this report multiple methods including numerical (skewness and kurtosis), graphical (histograms, Q-Q plots, and box plots), and formal normality tests (Shapiro-Wilk) were used to determine if the data were normally or abnormally distributed.[143,145] The accumulation of all the tests were used to make the final determination of normality.

The variables of interest for this analysis were sex, body mass index (BMI), age, neck muscle fatty infiltrate (MFI) measured from dual-echo fat/water (Dixon) magnetic resonance imaging (MRI) scans, numeric pain rating scale (NPRS), NDI score, sleep item response on the NDI (sleep NDI), and the sleep item response on the post-traumatic disability scale (sleep PDS). Sex, BMI, and age were collected at baseline; within 1 week of the MVC. For the variables MFI, NPRS, NDI, and both sleep item responses, there were data for each of the 4 time points of the

collection process: within 1 week (T1) post-motor vehicle collision (MVC), 2 week (T2) post-MVC, and both 3 months (T3) and 12 months (T4) post-MVC.

Sex was collected on the intake form with the options of male and female. It is recognized that the sex of a person is often not simply defined as male and female.[146] However, for the purpose of this study, it was determined that the contemporary designations of male and female would be used as the determinants of a person's sex-at-birth.[17,59,147] Because the responses were categorical and could not be distributed normally, it was not included as part of the data screening.

4.2.1 Data Cleaning

Prior to analyzing the data in SPSS, missing records were identified and managed in Microsoft Excel 2018 (Redmond, WA). Each participants information was entered in the spreadsheet and every variable was assessed for missing data by looking for empty cells within the table. For any variable recorded at baseline, if data were missing then the average of that variable based on group membership was used in that space. For missing data related to T2, T3, or T4, the data in the previous cell would be carried forward and used for all empty cells.

For the variables sex, age, and BMI there were no missing data points. For NDI, the only missing data were at T4; 8 participants in the persistent group and 11 in the recovered group. Neck MFI in the persistent group had 3 participants with no data available at any time (they were either claustrophobic and could not undergo the MRI scan, or other institutional policies prohibited them from having an MRI) along with 6 participants without available data at T4 as they were lost to attrition. For neck MFI in the recovered group, 1 participant had no data available due to claustrophobia and 10 participants were missing data at T4 again due to attrition. In recording the NPRS for the persistent group, 1 participant was missing data at T1, T3, and T4,

along with 7 other participants without data at T4, and in the recovered group 1 participant had no data at any time, and 10 others lacked data at T4. Finally, sleep NDI and sleep PDS responses were identical for both groups regarding missing data. For these variables, missing data was only an issue at T4, the persistent group had 8 participants along with 10 participants in the recovered group. For all variables, participants who had no data at T4 were lost to follow up.

For the group of participants that developed persistent symptoms, the 9 people who did not complete all the MRI data collection were compared to the 51 that did complete all data points to assess if there were any differences in their age or BMI. An independent *t*-test was used to evaluate for any differences for these two variables; for age there was no difference noted between those that completed all data points and those that did not, but the BMI was significantly higher in the group that did not have the MRI when compared to the group that did complete the scan ($p = 0.024$). For the recovered group, the 11 participants that did not complete all MRI data points were compared to the 26 that did to determine if a difference existed between them for the variables BMI and age. The independent *t*-test was used, and there were no significant differences between the groups on any variables.

In evaluating the difference in BMI between those that completed all the data points and those that did not, it was discovered that of the participants in the non-completion group, one had a BMI of 39.7 and another had 34.8. Of all 97 participants in the study, the BMI of 39.7 was the highest value recorded and the 34.8 was the 5^{th} highest value. Therefore, although there was a difference between the BMI of the people that were able to provide data points at all time points and those that did not in the persistent group, the two elevated scores may have skewed the data.

4.2.2 Numerical Methods

Skewness and kurtosis were evaluated for each of the variables; for the variables with four data points each was evaluated individually. There is no gold standard for what constitutes skewed or kurtotic data.[143,148] For determining if the data are skewed or kurtotic, the output of the SPSS analysis was used. In the output, total skewness and kurtosis were reported as a numerical value and if the range was within -1 to 1, the data were considered to have a normal distribution.[143]

Using the standard error is another option for assessing the numerical evaluation of skewness and kurtosis.[143,148] However, when evaluated across time and groups, the data would likely be bimodal with numerous high and low data points and not many in the middle. Thus, the distribution will create a large standard error and numerous variables will be deemed skewed or kurtotic. Because of this distribution pattern, the method of using the standard error for evaluating skewness and kurtosis was not used.

Using the range of -1 to 1 for evaluating skewness, 4/23 variables were found to be skewed: NDI at T3 and T4, sleep PDS T4, and MFI T1. In looking at these variables, it is recognized that 3 of the 4 were present at T3 and T4, which by this time point the differences between the groups, persistent symptoms and recovered, had been well established and thus the data were expected to be non-normally distributed.

Kurtosis was evaluated the same as skewness using the -1 to 1 scale, and in the SPSS output 4/23 variables were outside this range. The variables were MFI T1 and T2, NDI T4, and sleep PDS at T4. Similar to the skewness, 2 of the 4 variables were related to the most distal time points and would thus be expected to not be distributed normally. Only MFI at T1 and T2 unexpectedly were found to be kurtotic, but as it was not seen at T3 and T4 and since regression

is robust to minor deviations of normality, it was determined to retain the data in a raw (non-transformed) format.

4.2.3 Graphical Methods

Q-Q plots were used to evaluate the normal distribution of each variable. Most of the variables showed a proper distribution along the primary diagonal line with the exceptions being NDI at T3 and T4, and MFI at T1 and T2. The results for the NDI were expected based on the skewed data previously described and the groupings of the participants at these times. For the MFI, although it was found to be abnormally distributed at T1 and T2, it was not seen at all 4 time points: thus, it does not appear to be a systemic issue related to MFI.

In addition to the Q-Q plots, histograms of the spread of the data were examined to assess normality. The histogram for the variable NDI at T3 and T4 showed that the data were not normally distributed, which agrees with the Q-Q plot. However, MFI at T1 and T2 showed a normal distribution with the histogram, which is in opposition to the Q-Q plot, but with the Q-Q plot the deviation from normal appeared to be minimal. Histograms for the variables at T3 and T4 NPRS, sleep NDI, and sleep PDS showed a non-normal distribution. However, like the NDI at T3 and T4, these results would be expected based on the widely acknowledged assumption that 50% of those injured should recover within the first 3 months, as well, given these variables are ordinal level data, a non-normal distribution can be expected.[149] All other variables with the histogram showed a normal distribution.

4.2.4 Formal Normality Tests

The Shapiro-Wilk test is considered the most powerful normality test[145] and was used in the analysis of this data set. However, it is extremely sensitive to minor departures from a normal

distribution:[143] its results should be considered in combination with the previous assessments of normality.

For the Shapiro-Wilk test, the alpha level of 0.05 was used, and the null hypothesis was that the data were normally distributed: therefore, a probability lower than alpha were considered not normally distributed. For all the variables, only the NDI at T1 and NPRS at T1 were above the 0.05 level of significance with levels of 0.547 and 0.059 respectively. Thus, all the other variables, per the Shapiro-Wilk test, were considered non-normally distributed.

4.2.5 Outliers

To determine if there were any outliers in the data, box plots were created, displayed, and assessed. For each variable, the interquartile range (IQR) was reported as the quantity of the 25th and 75th percentile of the total score. The IQR was multiplied by 1.5 and added to the 75th percentile and subtracted from the 25th percentile to determine the highest and lowest scores that are not deemed to be an outlier. Any number above or below the 1.5 IQR is considered an outlier: if it is between the 1.5 and 3 IQR it is a mild outlier, whereas extreme outliers are reported if it is above 3.[143]

In the output, there were several outliers, all of them were above the 75th percentile, and none were below the 25th percentile: they were all considered mild outliers as none were beyond the 3 IQR. The outlier variables were BMI (1 participant), NDI at T3 (2 participants), sleep PDS at T4 (4 participants), and MFI at each time point had at least 3 outliers (with 5 distinct participants across all points). To evaluate the potential causes for the outliers, all outlier scores were assessed for accuracy by ensuring the data were accurately recorded in the spreadsheet and no typographical errors were found in the input. After reviewing the original data, it was determined that no typographical errors were present related to the outliers.

To assess the influence the outliers may have had on the overall normality, the outliers were transformed using the Winsor method, which takes their current value and adjusts it to the 1.5 IQR above the 75th percentile so that it is at the outermost portion of the whisker of the box plot but it is no longer considered an outlier.[150] The statistical tests were run again using the Winsorized variables and all numerical, graphical, and normality test results were compared to the original output.

For NDI and sleep PDS responses it was expected that there were outliers at T3 and T4. By these time points, many participants are dichotomized into the recovered or persistent groups and the data were shifted towards a non-normal distribution, which is different than the earlier times when the data were normally distributed. However, if someone still had severe symptoms at T3 or T4, they may be outside the 1.5 IQR, whereas at earlier time points when the data were evenly distributed those same scores were within the 1.5 IQR. Use of the Winsor method for NDI and sleep PDS responses indicated there were no changes in the skewness or kurtosis: however, the Q-Q plot for NDI at T3 was now normally distributed, and all other graphical evaluations were unchanged when compared to the untransformed data, as well, there was no change in the Shapiro-Wilk test.

For BMI and MFI, the variables are not based upon symptoms, and it is unknown if the data were affecting the overall normality. Using the Winsor method, there were no changes in any of the normality tests for BMI. For MFI assessing the Q-Q plot at T2, the data were normally distributed, but all other graphical plots were unchanged, and the Shapiro-Wilk test was unchanged.

With all these tests taken into consideration, there are a few variables that are not normally distributed, primarily MFI at T1 and T2. However, given that MFI does not

demonstrate this distribution at T3 and T4, and the majority of the normality tests and graphs do not demonstrate significant variations from normal, it is safe to state the data were normally distributed and parametric tests could be used in the analysis. For the other variables, based on the expected recovery rates of people following an MVC and the bimodal distribution of the data, it is was expected that some would be non-normally distributed. In this study, the variables that demonstrated a non-normal distribution were centered around T3 and T4 and included NDI, NPRS, sleep NDI, and sleep PDS. As earlier time points demonstrated normal distribution, there should be no concerns with using parametric tests during the statistical analysis.

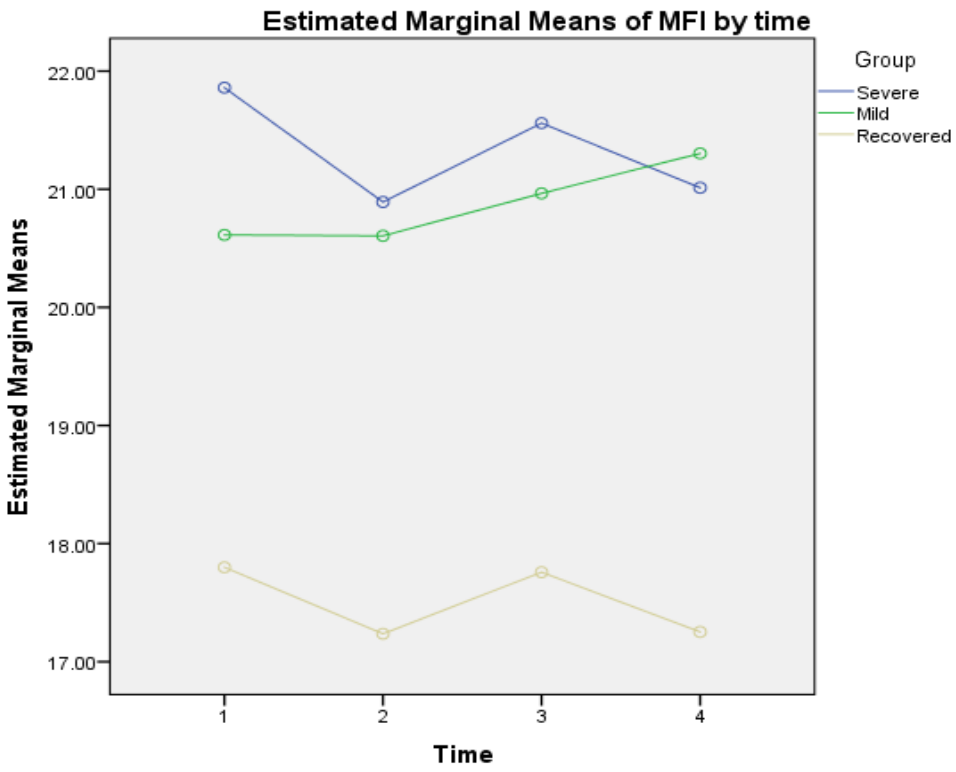
4.3 Specific Aim #1

The primary aim for this study was to evaluate the relationship between the independent variables (IV): MFI, sex, age, BMI, NPRS, NDI, sleep NDI, and sleep PDS responses on the dependent variable (DV) group membership that consisted of persistent symptoms and recovered amongst participants following an MVC. Group membership was based on total NDI scores at 12 months and was separated into two groups: 1) persistent symptoms in anyone with an NDI 10% and above, and 2) recovered as any person with an NDI score below 10%.

In this study, participants were divided into two groups (persistent and recovered) instead of the three group model previously described in the literature (recovered, mild, and severe).[4] In the three-group model, which is also based on NDI scores, the recovered group is 0 to 8%, mild is classified from 10 to 28%, and moderate to severe is greater than 30%. The idea to use two groups instead of three was made for several reasons. The notation of persistent and recovered is more clinically applicable to those suffering with whiplash associated disorder (WAD). It may be that people that are coping with mild symptoms could (although there is the possibility that they may not) benefit from treatment regardless if their symptoms are not as debilitating as those who

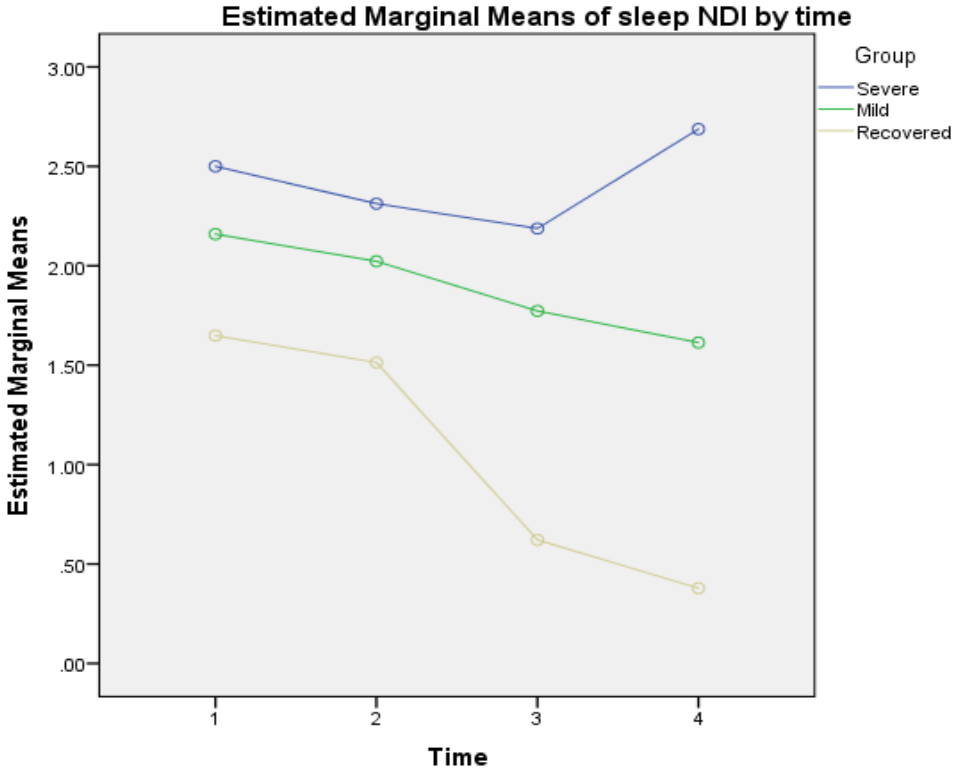
are classified as moderate to severe. A second reason had to do with the large minimum detectable change (MDC) reported for the NDI; the MDC is the smallest change that must occur between two related scores to exceed the threshold of error.[136] The MDC for the NDI has been reported to be as high as 19%,[137] which essentially means that someone with a score of 12% could be in the same category as someone with 30% impairment. Therefore, using a cutoff of 30% could inadvertently misrepresent the total number of people who could benefit from treatment. Finally, in evaluating the variables across the four time points in this study, when the variables were analyzed using the three-group model, there were several variables where the mild and moderate to severe categories resembled each other rather than demonstrating distinctive groups. However, the mild and severe groups appeared to be different when compared to the recovered group. See Figures 4.1-4.3 for the profile plots of the variables across time.

Figure 4.1 MFI across time by group



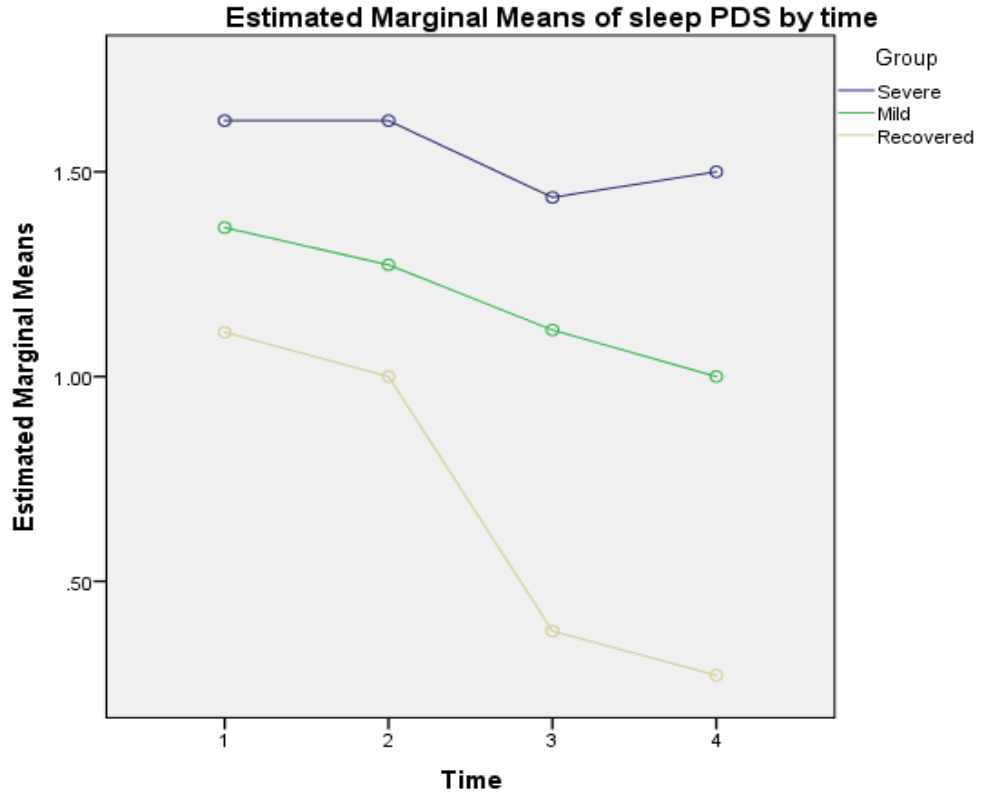
Abbreviations: MFI, Muscle Fatty Infiltrate.

Figure 4.2 Sleep NDI across time by group



Abbreviations: Sleep NDI; Sleep response on Neck Disability Index.

Figure 4.3 Sleep PDS across time by group



Abbreviations: Sleep PDS; Sleep response on Post-traumatic Diagnostic Scale

4.3.1 Baseline Demographics

Prior to evaluating the relationship between the IVs and the DV, the baseline demographics needed to be established, in addition, it needed to be determined if there were any differences between the groups at this time. To assess the variables, a series of tests were performed to establish the means for the groups and determine if any statistically significant differences between the groups were present. For the variables with interval or ratio level data (age, BMI, MFI, and NDI) an independent *t*-test was used,[136] for ordinal level data (NPRS, sleep NDI, and sleep PDS) a Mann-Whitney U-test was used,[136] and for categorical data (sex) a Chi-Square was used.[136]

For the independent *t*-test, with the four variables evaluated, only MFI (p = 0.007) and NDI (p = 0.004) were found to be statistically different between the groups. See Table 4.1 for means and p-values for all interval and ratio level variables.

Table 4.1: Baseline characteristics and differences for continuous data (independent *t*-test)

Variable	**Means of recovered group (n=37)**	**Means of persistent symptoms group (n=60)**	**p-value**
Age (years)	34.1	35.0	0.68
BMI	25.27	25.15	0.90
MFI T1	17.80	20.95	0.007*
NDI T1	30.0	39.9	0.004*

*Abbreviations: BMI, Body Mass Index; MFI, Muscle Fatty Infiltrate; NDI, Neck Disability Index; * Reached statistical significance with alpha at 0.05.*

A Mann-Whitney U test was used to evaluate differences between the groups for the variables with ordinal level data. Of the three variables assessed at baseline only sleep NDI was found to be significantly different between the groups (p = 0.027). See Table 4.2 for median responses and significance level of each variable.

Table 4.2: Baseline characteristics and differences for ordinal data (Mann-Whitney U test)

Variable	Median of recovered group (n=37)	Median of persistent symptoms group (n=60)	Asymp. Sig (2-tailed)
NPRS T1	5.0	5.0	0.20
Sleep NDI T1	1.0	2.0	0.027*
Sleep PDS T1	1.0	2.0	0.12

*Abbreviations: NPRS, Numeric Pain Rating Scale; Sleep NDI, Sleep response on Neck Disability Index; Sleep PDS, Sleep response on Post-traumatic Distress Scale; * Reached statistical significance with alpha at 0.05.*

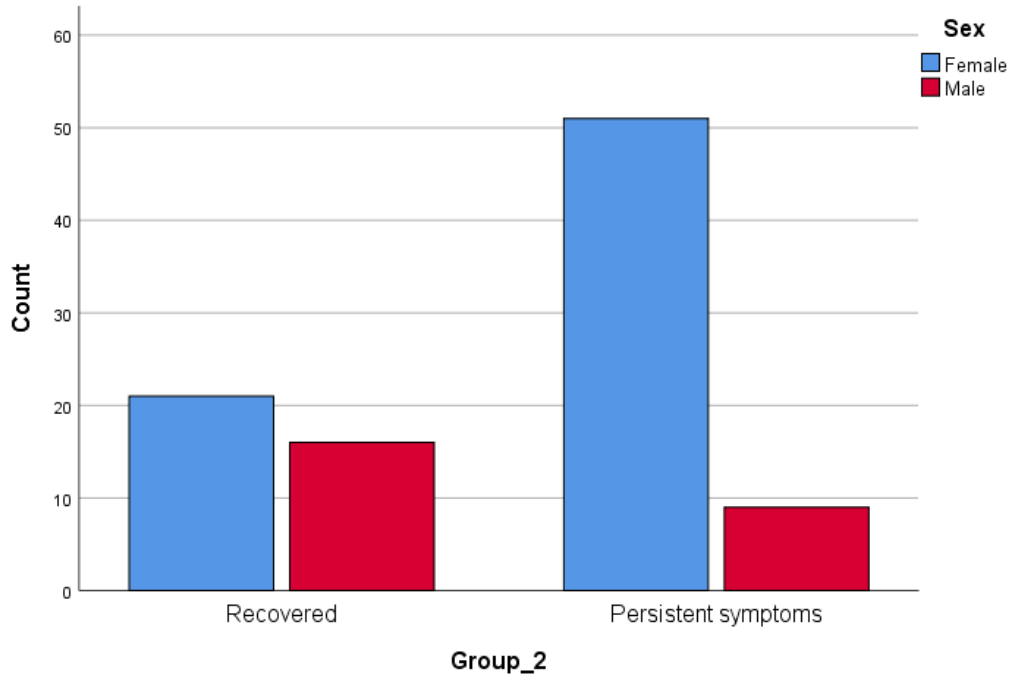
A Chi-Square evaluated the relative proportions of males to females in the groups. Using this test and the Pearson Chi-Square statistic, there were proportionally more females in the persistent group (85%) than in the recovered group (57%), $\chi^2 = 9.54$, $p = 0.002$. See Table 4.3 for the results from the Chi square, and Figure 4.4 for a graphical representation of differences between the groups.

Table 4.3: Baseline characteristics for nominal data (Chi-Square test)

Variable	Recovered Female/Male (n=37)	Persistent symptoms Female/Male (n=60)	Pearson Chi-Square statistic	p - value
Sex	21/16	51/9	9.54	0.002*

** Reached statistical significance with alpha at 0.05.*

Figure 4.4 Sex distribution between recovered and persistent symptoms



The ordinal data was assessed for differences at baseline using non-parametric tests, however, when using logistic regression the variables need to be classified as either categorical or continuous prior to placing them in the model.[144] Therefore, prior to running any analyses, the variables sleep NDI, sleep PDS, and the NPRS were examined to determine how to categorize them for the model. The sleep NDI and sleep PDS were examined in a similar manner as their scales were shorter, 0 to 5, and 0 to 3, respectively when compared to the NPRS which has a scale of 0 to 10.

The sleep NDI and sleep PDS were initially examined as continuous variables by creating separate variables for each score on their scale (0 to 5 for NDI and 0 to 3 for PDS) and dummy codes to represent all but the single score that was evaluated. Prior to this evaluation, examining the contingency table showed that there were not enough participants in the sleep NDI group

who reported a score of 5. Therefore, the scores of 4 and 5 were combined into one variable, and the final groupings for the sleep NDI were 0, 1, 2, 3, and 4/5. The sleep PDS scores did not have this issue and the groups were 0, 1, 2, and 3. Once the variables were created, individual logistic regression models were created for the sleep NDI and sleep PDS using the enter method with all variables placed in the model. The unstandardized beta coefficients were examined to determine if they were different across the responses. In both models, none of the scores demonstrated a statistically significant difference when compared to one another.

The sleep NDI and sleep PDS were then examined as categorical variables. While evaluating the scores as a continuous variable it was noted that for both the sleep NDI and sleep PDS the response of 0 showed a negative beta value, which means that with that response there is a protection from developing persistent symptoms. On the other hand, all other scores reported positive beta values meaning that they were predictive of developing persistent symptoms. Therefore, the categorical cutoff for both the sleep NDI and sleep PDS were a score of 0 for one group, and all other scores (NDI 1 to 5, PDS 1 to 3) for the other group. Using these categories, separate logistic regression models were created using the enter method with the categorical variables described for sleep NDI and sleep PDS responses at T1. Both models demonstrated that the groups that had responses that were non-zero were predictors of persistent symptoms with positive beta values.

The models were then compared using the likelihood ratio test to determine if there was a statistical difference between them. For both the sleep NDI and sleep PDS, the likelihood ratio test was found to be non-significant, which means that there was no difference in the way that the models predicted persistent symptoms; using categorical or continuous data gave the same results. Therefore, since the variables did not demonstrate an increasingly larger beta value

across response scores, and the overall models were not significantly different from each other, it was determined that using the sleep scores as categorical variables was the best option for the creation of the final model.

The NPRS is scored on a scale from 0 to 10, and it was initially examined on the contingency table and it was recognized that the data was spread much like a bell curve with few scores at either end and approximately 50% of the scores at 4 through 6. Therefore, each individual score could not be evaluated as it was for the sleep responses, and a single logistic regression model using the enter method and the NPRS scores as continuous with all scores 0 to 10 was created. The results in the model showed a beta value that was positive, meaning that for increasing scores on the NPRS there was a relationship with an increased risk for developing persistent symptoms.

From a categorical perspective, the NPRS was then broken into two groups, 0 to 5 for one group and 6 to 10 for the other which was based on previous research that found that 5.5 out of 10 was the point which seemed to differentiate those who would progress to develop chronic symptoms.[17] Using this variable, another logistic regression model was created and found to have a significant positive beta value for the group with higher scores indicating a relationship with persistent symptoms.

When the two models were compared with the likelihood ratio test, the difference between them was not found to be significant meaning that the variables were similar when placed in the model. However, in opposition to the sleep responses, for the NPRS it was determined that since it has a larger scale 0 to 10 and the data is normally distributed across the scores that using it as a continuous variable would be appropriate for the final model.

The final evaluation procedure prior to building the logistic regression model for the baseline data was to determine whether any multicollinearity existed between any of the IVs. Because some of the variables are at the interval level and others are at the ordinal level, a Spearman's Rho correlation was used; because sex was defined as a categorical variable it was not part of the analysis. Although the sleep response scores from the NDI and PDS were used as categorical data in the model, their individual data points could be used as part of the correlation matrix to help determine if multicollinearity existed.

Prior reports have defined multicollinearity as being present when the correlation coefficient is in the range of 0.70 to 0.80.[143] A conservative approach is to recognize multicollinearity when the correlation coefficient is above 0.75 as it is believed that 0.80 is too strict of a cut off.[143] If the cut off is too high, there is the potential to keep variables in the model that are essentially represent the same influence on the DV. With this assessment, the only variables in the correlation matrix that approached the 0.75 level was sleep NDI and sleep PDS (0.68). See Table 4.4 for all coefficient values.

Using two variables that measure the same construct likely will be associated with multicollinearity.[143] Since the r value for the variables did not reach the 0.75 threshold (0.68), to further evaluate the potential for multicollinearity between sleep NDI and sleep PDS the relationship between them was evaluated at each of the four time points. A Spearman's Rho was performed comparing sleep NDI T1 to sleep PDS T1 ($r = 0.68$), sleep NDI T2 to sleep PDS T2 ($r = 0.66$), sleep NDI T3 to sleep PDS T3 ($r = 0.76$), and sleep NDI T4 to sleep PDS T4 ($r = 0.72$). With these correlation coefficients, it appears that multicollinearity is only present at T3. However, as sleep NDI and sleep PDS evaluate the same construct, sleep disturbance, albeit in a slightly different manner, the NDI uses a 0 to 5 scale and the PDS uses a 0 to 3 scale, and

because they demonstrate multicollinearity at T3, it was deemed inappropriate to keep both variables in the final model. Given that the NDI scale is more readily available to practitioners and researchers, as well as its possible greater significance in this population based on its difference with the Mann-Whitney U tests at baseline, it was determined to keep the sleep NDI scale for inclusion in the model and discard the sleep PDS responses.

Table 4.4 Multi-collinearity correlation coefficient of all IVs

	BMI	**Age**	**MFI T1**	**NPRS T1**	**Sleep NDI T1**
Age	0.24				
MFI T1	0.17	0.41			
NPRS T1	0.11	-0.05	0.001		
Sleep NDI T1	-0.02	-0.03	0.25	0.47	
Sleep PDS T1	-0.05	-0.19	0.07	0.27	0.68

Abbreviations: BMI, Body Mass Index; MFI, Muscle Fatty Infiltrate; NPRS, Numeric Pain Rating Scale, Sleep NDI, Sleep response on Neck Disability Index; Sleep PDS, Sleep response on Post-traumatic Distress Scale.

4.3.2 Logistic Regression at Baseline

As the overall theme of the study was exploratory, it was deemed preferential to err on the side of a Type I error (a false positive) rather than a Type II error (a false negative).[136] If a Type I error were to be present then people may receive treatment that may not be justified, whereas a Type II error could create a scenario where those that may benefit from treatment may not receive appropriate care and could be harmed or at least have their suffering prolonged. Therefore, in order to minimize the risk of a Type II error, the alpha level of 0.10 was used instead of the standard 0.05 for inclusion in the model to ensure that all variables of significance,

or even potential significance, were included. As well, using a higher p-value may allow the entrance of some variables that may not be significant at the 0.05 level at baseline, but may reach that level over time.

In addition to the increased alpha value, a backward conditional method for entering values into the logistic regression model was used; using a forward conditional method, variables require a more stringent criteria for being placed into the model.[143] By allowing more variables into the model the likelihood of rejecting a variable (a Type II error) is reduced.

For the logistic regression at baseline, the outcome of interest was to determine which IVs were predictive of the participants transition to persistent symptoms based on their NDI scores at 12-months post-MVC. All variables were put into the model and those at the 0.10 level of significance or below were kept for evaluation at the other three time points.

However, prior to building the model, it was determined that using the NDI as a predictor of the DV was inappropriate because this analysis was examining the variables and the change in strength of the relationship with group membership over time. Given that the score of the NDI at 12 months was used as the determining factor in group membership, it would be expected that the model at T4 would include only the NDI score, with 100% accuracy, thus rendering that portion of the analysis, and the influence of any other variables useless.

Using the backward conditional method to create the initial model, the IVs at T1 BMI, age, and NPRS all were all found to be non-significant as predictors of the DV. See Table 4.5 for significance values at time of removal. The main effects for IVs at T1 sex, MFI, and sleep NDI were all found to have significance values below 0.10 and would be included in the model and evaluated to assess their change over time. See Table 4.6 for significance values, unstandardized

beta values, OR and 95% CI for each variable. Within this model using these three variables, the pseudo R^2 was calculated using the Nagelkerke R^2 and reported as 0.27. The Nagelkerke R^2 was used instead of the Cox and Snell R^2 as it is reported on a 0 to 1.0 scale which corresponds more directly to a true R^2, whereas the Cox and Snell reports on a scale from 0 to 0.75, thus the results will always be lower and are more difficult to interpret and compare.[143] For being able to classify participants based on the variables in this model, an overall percentage of 76.3% were classified correctly with 85.0% correct for persistent symptoms and 62.2% correct for the recovered group. These results are held in relation to the constant model that states without any variables in the model that if all participants were classified as having persistent symptoms, the correct prediction would be reported 61.9% of the time.

Table 4.5 Significance values for variables not included in final model at T1

Variables	**Significance at point of removal**
NPRS	0.27
BMI	0.49
Age	0.40

Abbreviations: Sleep NDI, Sleep response on Neck Disability Index; BMI, Body Mass Index.

Table 4.6 p-value, unstandardized beta values, odds ratio, and confidence interval for main effect variables kept in model at T1.

Variable	p-value	Unstandardized beta value	95% CI for beta value	Odds ratio	95% CI for odds ratio
Sex (female)	0.013	1.29	0.27 to 2.31	3.63	1.31 to 10.05
MFI	0.066	0.10	0.0 to 0.20	1.10	0.99 to 1.22
Sleep NDI	0.016	1.48	0.28 to 2.67	4.38	1.32 to 14.51

Abbreviations: MFI, Muscle Fatty Infiltrate; NPRS, Numeric Pain Rating Scale; CI, Confidence Interval.

Before the variables at T2 could be analyzed to evaluate the changes in magnitude of beta and OR, it needed to be recognized that the main effects may not fully explain the relationships between the variables and group membership. As well, it is possible that some of the variables that were found to be non-significant could have an influence when part of an interaction. Therefore, interaction effects between variables were evaluated in conjunction with the main effects that were previously assessed.

Determining which variables to include as part of the interaction effects was difficult as the primary aim of the research was exploratory, and all the variables could theoretically be included. Within this study, all variables have been found in previous research to have at least some prognostic validity in determining persistent symptoms in people with WAD.[17] However, in an attempt to be as objective as possible, and to keep the number of interactions to a

reasonable amount, it was determined that only the variables that had a statistically significant difference between the groups, persistent symptoms and recovered, at baseline would be included as interactions. Therefore, the interactions included in this analysis were as follows: sex by MFI; sex by sleep NDI; and MFI by sleep NDI.

Running a backward conditional method with the all IVs (sex, BMI, age, MFI, NPRS, and sleep NDI) as well as interactions (sex by MFI, sex by sleep NDI, and MFI by sleep NDI) the final predictors in the model for determining group membership were: sex, MFI, sleep NDI, and no interactions were found to be significant. When compared to the model with only the main effects, there was no difference between the two models as all the variables in the final model were the same and their unstandardized beta values were identical as well. In addition, the Nagelkerke R^2 of 0.27 and the classification of groups (overall 76.3%, persistent group 85.0%, and recovered 62.2%) were all unchanged as well.

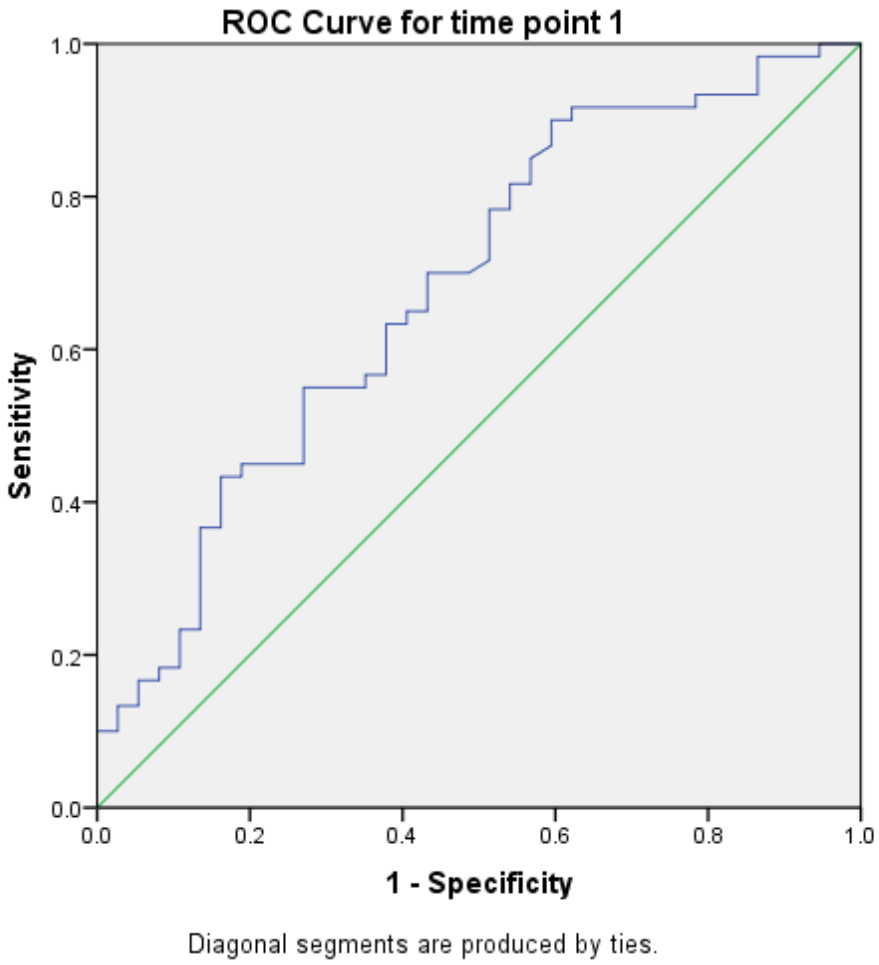
The percentages for correct prediction can also be reported as the sensitivity and specificity for the models as a way determine how accurate the predictions are for who will develop long-term symptoms and who will recover. The sensitivity of the test is based on the presence or absence of the variables in those with persistent symptoms and will provide evidence for a Type II error (a false negative). In this model, the sensitivity is 85%; this means that there is a 15% chance for a false negative. Thus, if someone does not have the variables present, (female sex, no sleep disturbances, and low MFI) it is likely they will have a true negative and fall into the recovered group. Specificity is based on presence or absence of the variables in those that are found to be recovered. In this model, the specificity is 62.2%, which means there is a 38.8% chance of a false positive, however, this is only somewhat better than chance. As the possibility

of a false positive hovers around 40%, it is recognized that the presence of the variables is not as strong an indicator for ruling someone into the persistent group.

Minimizing a Type II error was a focus of this project, as a false negative can be problematic for those that develop persistent symptoms as they may end up not receiving treatment which could benefit them. Therefore, being able to discriminate between the two groups is of the utmost importance. To assess the overall discrimination of the proposed model, a ROC curve was generated along with its AUC calculation.

The purpose of the ROC curve and the AUC is to determine how well the variables help to discriminate between the two possible outcomes:[143] recovered and persistent symptoms. The ROC curve is created by using a combination of true-positive and false-positive rates to predict the desired outcome.[143] The ROC curve is a graphical representation of these rates, whereas the AUC, which ranges from 0.5 to 1.0, is the quantitative result of the ROC curve and can be used to state whether the discrimination of the variables is considered to have no discrimination (equal to 0.5), poor discrimination (greater than 0.5 and less than 0.7), acceptable discrimination (greater than or equal to 0.7 to less than 0.8), very good discrimination (greater than or equal to 0.8 to less than 0.9), or excellent discrimination (greater than or equal to 0.9) results based on its output.[143] See Figure 4.5 for ROC curve for the baseline variables used in the model. Using this data, the AUC was reported at 0.68, which is considered poor discrimination, and it had a 95% CI of 0.57 to 0.79 with a significance value of 0.003; these CIs ranged from poor to acceptable discrimination.

Figure 4.5 ROC curve for variables at time point 1



Abbreviations: ROC; Receiver Operator Characteristic

The two models previously evaluated, those with interaction effects and those without them, provided similar outputs. Even though no interaction effects were found to be significant predictors of persistent symptoms at baseline, there is no guarantee that they will remain non-significant over time as variables generally do not function in isolation. Therefore, in analyzing the output for the subsequent 3 time points, using the model with all the interaction effects was evaluated.

4.4 Specific Aim #2

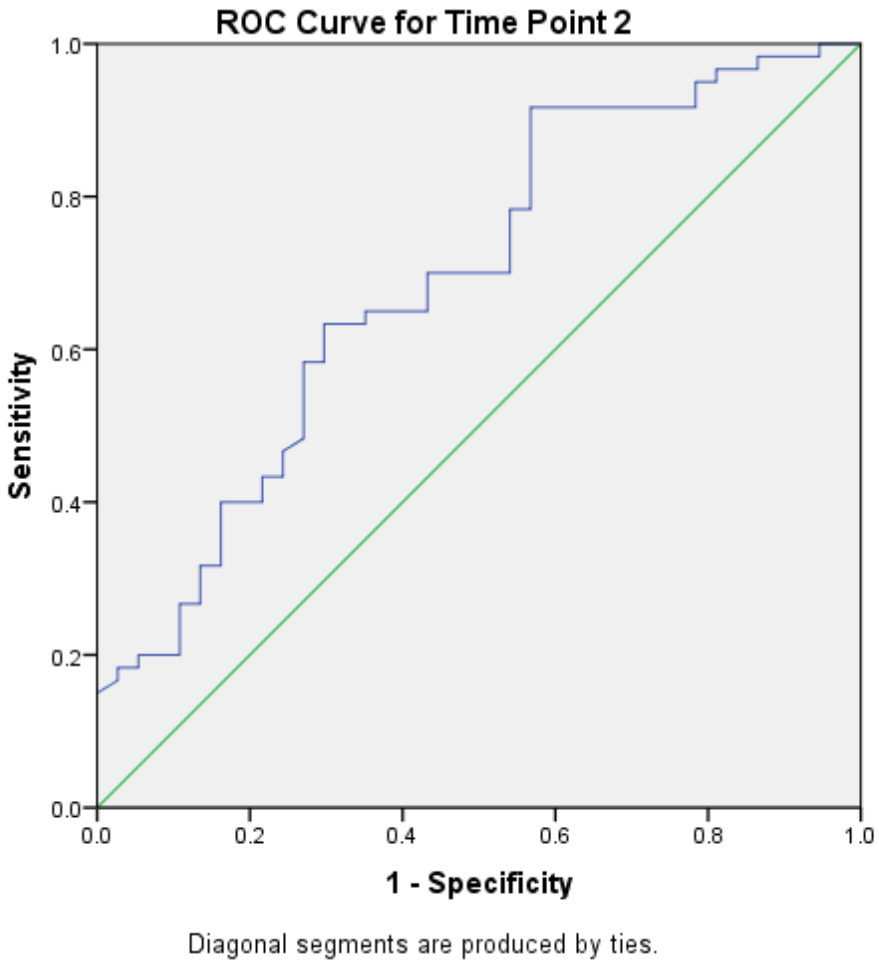
The second aim of this study was to determine if the predictive strength of the IVs (beta values, OR, 95% CI, and pseudo R^2) reported in section 4.3.2 and their relationship with persistent symptoms associated with WAD changed in magnitude or direction at specified times over a 12 month period following the MVC.

At T2, T3, T4, the variables reported in the final model at baseline were used with the same statistical analysis previously described for the logistic regression to determine changes in magnitude or direction of the IVs prediction over time. A backward stepwise conditional logistic regression was used with an alpha level of 0.10, and all the IVs included: sex, MFI, sleep NDI, and the interactions of sex by MFI, sex by sleep NDI, and MFI by sleep NDI.

4.4.1 Logistic Regression at Time Point #2

T2 was defined as 2 weeks post-MVC. From this analysis, two predictors were determined to have statistical significance in terms of predicting group membership: sex and MFI. With these variables the Nagelkerke R^2 was reported to be 0.22. The classification of persistent symptoms and recovered reported a prediction of 70.1% correct overall, and it correctly identified 86.7% (sensitivity) of those with persistent symptoms and 43.2% (specificity) of recovered participants. See Figure 4.6 for the ROC curve for T2. Based on the ROC curve at T2, the AUC was a reported 0.69, which is a poor level of discrimination, a 95% CI of 0.58 to 0.80 which ranges from poor to very good, and it was statistically significant with a significance level of 0.002.

Figure 4.6 ROC curve for variables at time point 2



Abbreviations: ROC; Receiver Operator Characteristic

For the individual variables, sex (female), the p-value was 0.01, and an unstandardized beta value at 1.31, and OR of 3.71, with a 95% CI of 1.37 to 10.07. MFI had a p-value of 0.021, an unstandardized beta value of 0.13, OR of 1.14, and 95% CI of 1.03 to 1.26. Sleep NDI had a p-value of 0.14 and was thus not included in the final model at T2, and none of the interactions were found to be significant predictors either. See Table 4.7 for results from the variables included in the model. The variables not included in the model are reported in Table 4.8 along with their significance level at their point of exclusion.

Table 4.7: p-value, unstandardized beta values, odds ratio, and confidence interval for variables kept in model at T2

Variable	p-value	Unstandardized beta value	95% CI for beta value	Odds ratio	95% CI for odds ratio
Sex (female)	0.008	1.31	0.31 to 2.31	3.71	1.37 to 10.07
MFI	0.021	0.13	0.03 to 0.23	1.14	1.03 to 1.26

Abbreviations: MFI, Muscle Fatty Infiltrate; CI, Confidence Interval.

Table 4.8 Significance values for variables not included in final model at T2

Variables	Significance at point of removal
Sex by MFI	0.92
Sex by sleep NDI	0.47
MFI by sleep NDI	0.50
Sleep NDI	0.14

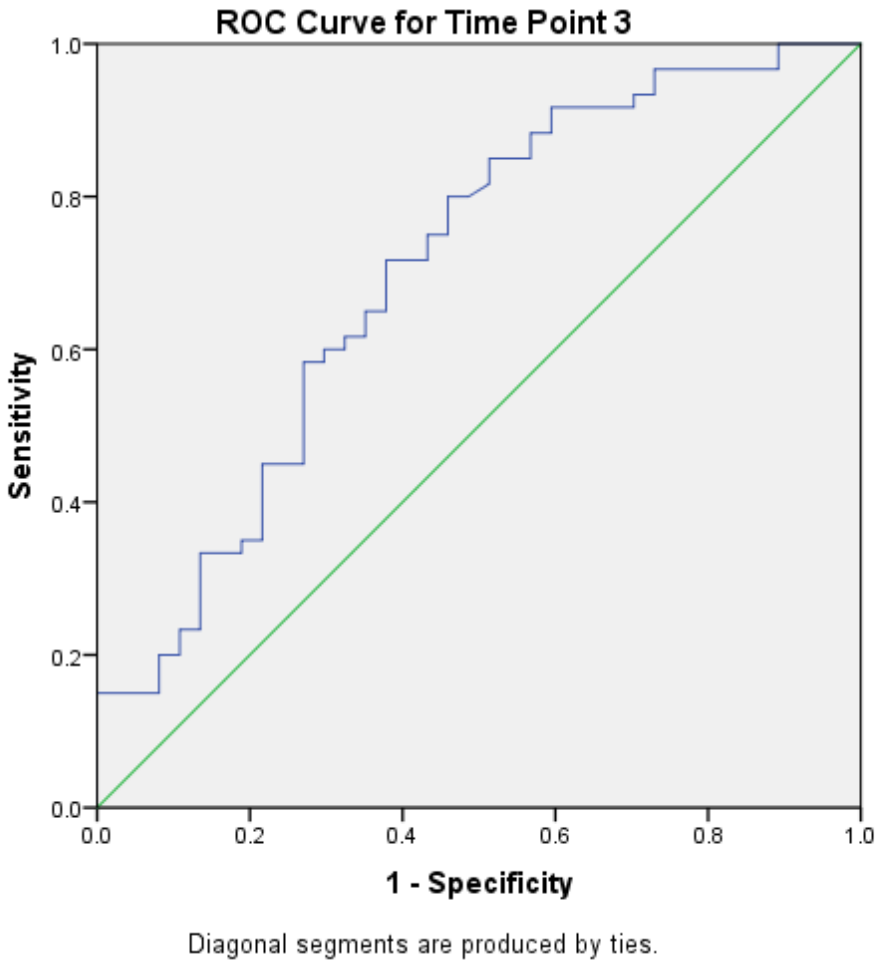
Abbreviations: MFI, Muscle Fatty Infiltrate; Sleep NDI, Sleep response on Neck Disability Index; NPRS, Numeric Pain Rating Scale.

4.4.2 Logistic Regression at Time Point #3

In this analysis, T3 was defined as 3 months post-MVC. The final variables included in the model at this time were: MFI, sex, and sleep NDI. With these variables, the Nagelkerke R^2 was 0.40. The correct classification of all participants using these variables was 79.4%, with 88.3% correct for persistent symptoms (sensitivity) and 64.9% for the recovered group (specificity). See Figure 4.7 for ROC curve at T3. Based on the ROC curve, the AUC was 0.70

which is classified as acceptable, a 95% CI from 0.59 to 0.81, and these numbers range from poor to very good, and it was a statistically significant finding with a significance value of 0.001.

Figure 4.7 ROC curve for variables at time point 3



Abbreviations: ROC; Receiver Operator Characteristic

For the variables at T3, MFI reported a p-value of 0.044, an unstandardized beta value of 0.11, OR of 1.11, and a 95% CI of 1.00 to 1.24. Sex reported a p-value of 0.009, an unstandardized beta value of 1.50, OR of 4.48, and 95% CI of 1.45 to 13.87. Finally, sleep NDI reported p-value of < 0.001, an unstandardized beta value of 2.05, OR of 7.75, and a 95% CI of

2.77 to 21.66. See Table 4.9 for the results of these predictors. See Table 4.10 for all variables excluded from the model and their significance values at the time of exclusion.

Table 4.9: p-value, unstandardized beta values, odds ratio, and confidence interval for variables kept in model at T3

Variable	p-value	Unstandardized beta value	95% CI for beta value	Odds ratio	95% CI for odds ratio
MFI	0.044	0.11	0.0 to 0.22	1.11	1.00 to 1.24
Sex (female)	0.009	1.50	0.37 to 2.63	4.48	1.45 to 13.87
Sleep NDI	< 0.001	2.05	1.02 to 3.08	7.75	2.77 to 21.66

Abbreviations: MFI, Muscle Fatty Infiltrate; NPRS, Numeric Pain Rating Scale; Sleep NDI, Sleep response on Neck Disability Index; CI, Confidence Interval.

Table 4.10 Significance values for variables not included in final model at T3

Variables	Significance at point of removal
Sex by MFI	0.77
MFI by sleep NDI	0.23
Sex by Sleep NDI	0.70

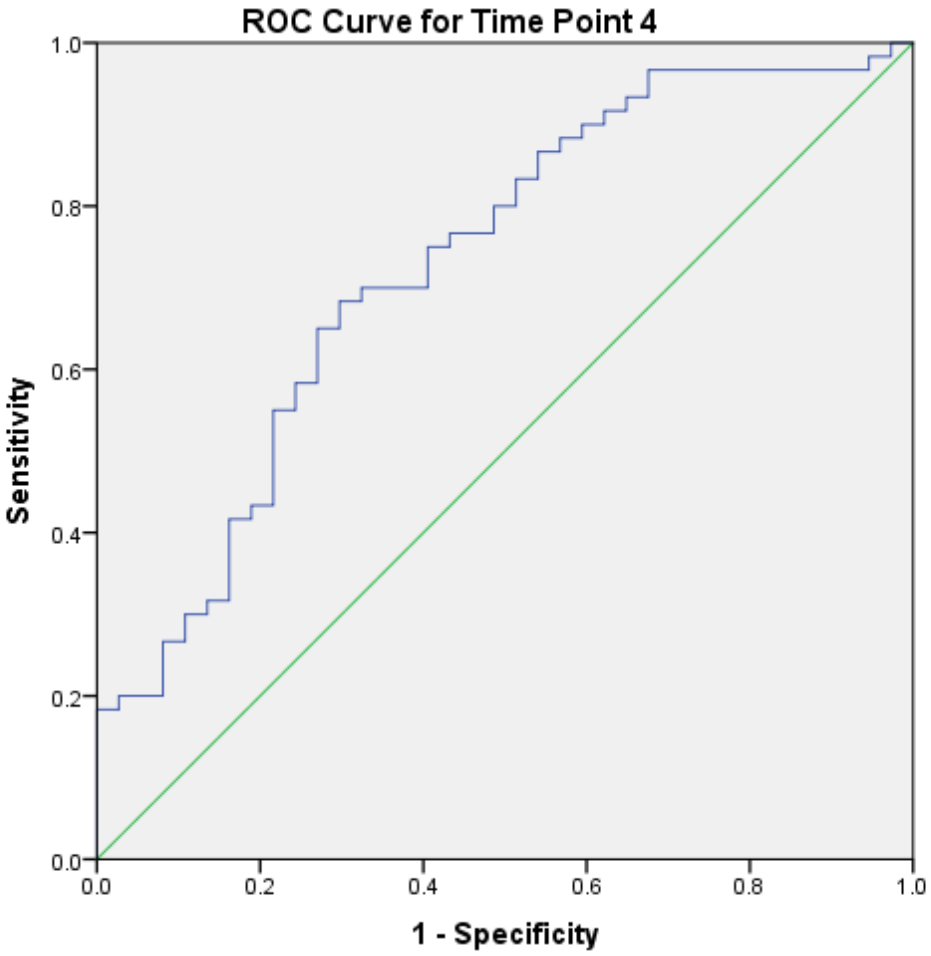
Abbreviations: MFI, Muscle Fatty Infiltrate; Sleep NDI, Sleep response on Neck Disability Index.

4.4.3 Logistic Regression at Time Point #4

For this analysis, T4 was defined as 12 months post-MVC. The variables included in the final model were sex, MFI, and sleep NDI. Based on these variables, the Nagelkerke R^2 was

reported as 0.56. The correct classification of all participants was 85.6%, with persistent symptoms at 86.7% (sensitivity) and recovered group at 83.8% (specificity). See Figure 4.8 for ROC curve at T4. From this ROC curve the AUC is reported as 0.73 which is considered acceptable discrimination, the 95% CI ranges from 0.62 to 0.83 and these numbers are deemed poor to very good, and the results are statistically significant with a significance value of < 0.001.

Figure 4.8 ROC curve for variables at time point 4



Abbreviations: ROC; Receiver Operator Characteristic

For each variable, sex had a p-value of 0.011, an unstandardized beta value of 1.69, OR of 5.42, and 95% CI of 1.48 to 19.82. MFI had a p-value of 0.027, an unstandardized beta value of 0.143, OR of 1.15, and 95% CI of 1.02 to 1.31. Finally, sleep NDI had a p-value of < 0.001, unstandardized beta value of 3.07, OR of 21.53, and 95% CI of 6.39 to 72.51. See Table 4.11 for the results of these predictors. See Table 4.12 for all the variables excluded from the model and their significance levels at the time of exclusion.

Table 4.11: p-value, unstandardized beta values, odds ratio, and confidence interval for variables kept in model at T4

Variable	**p-value**	**Unstandardized beta value**	**95% CI for beta value**	**Odds ratio**	**95% CI for odds ratio**
Sex (female)	0.011	1.69	0.39 to 2.99	5.42	1.48 to 19.82
MFI	0.027	0.14	0.02 to 0.27	1.15	1.02 to 1.31
Sleep NDI	< 0.001	3.07	1.85 to 4.28	21.53	6.39 to 72.51

Abbreviations: MFI, Muscle Fatty Infiltrate; NPRS, Numeric Pain Rating Scale; Sleep NDI, Sleep response on Neck Disability Index, CI, Confidence Interval.

Table 4.12 Significance values for variables not included in final model at T4

Variables	Significance at point of removal
Sex by MFI	0.77
MFI by sleep NDI	0.34
Sex by sleep NDI	0.77

Abbreviations: MFI, Muscle Fatty Infiltrate; Sleep NDI, Sleep response on Neck Disability Index.

To examine how the overall model at each time point compared with one another, see Table 4.13.

Table 4.13 Model summary across all four time points

	Nagelkerke R^2	Percentage correct overall	Percentage correct for persistent symptoms group (sensitivity)	Percentage correct for recovered group (specificity)
Time point 1	0.27	76.3	85.0	62.2
Time point 2	0.22	70.1	86.7	43.2
Time point 3	0.40	79.4	88.3	64.9
Time point 4	0.56	86.6	86.7	83.8

4.4.4 Comparisons Across Time

To assess changes in the variables across time, each variable's OR and 95% CI was evaluated in comparison to the previous time point. For a variable to be considered for this part of the analysis, it had to be significant in relation to the alpha level for prediction of persistent

symptoms for at least 3 of the 4 time points. If the variable was not significant at one of the time points, the OR and 95% CI that was reported at the previous time was used as a placeholder. This criterion was used to evaluate the variables sex, MFI, and sleep NDI response. MFI and sex were significant at all time points, and sleep NDI was significant at 3 of the 4 time points (it demonstrated non-significance at T2), with the lack of data at T2 for sleep NDI, the data collected at T1 was used for that portion of the analysis.

To determine if a significant change occurred regarding group membership prediction with either the OR or the CI, the metric was evaluated against the previous time point. To be considered statistically different from the previous time point, the OR must be beyond the range of the previous 95% CI; this should ensure a 95% chance that the two units are truly different. For the CIs, if there was no overlap between the two time points then they are said to be statistically different.

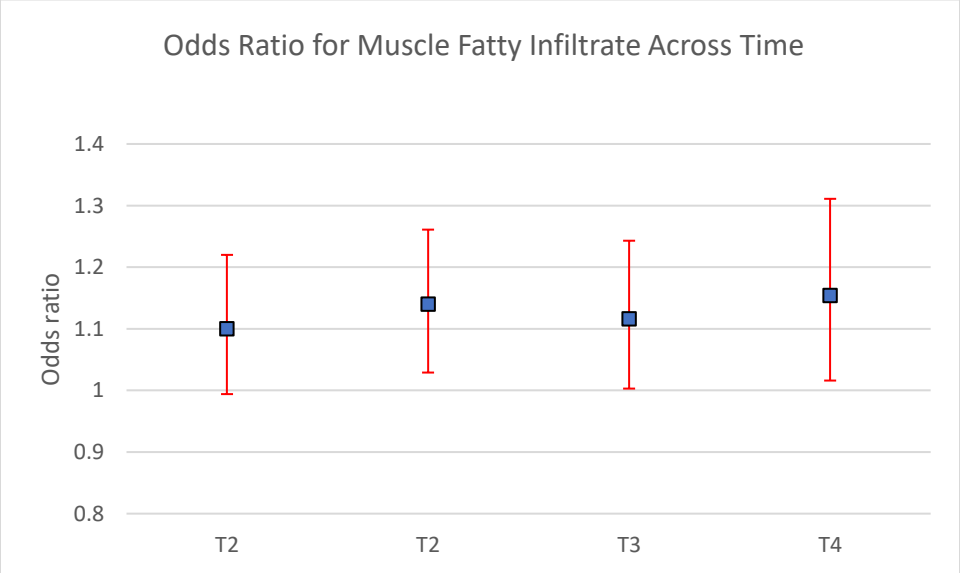
The first variable evaluated was MFI; see Table 4.14 for its OR and 95% CIs at each of the four time points and Figure 4.9 for the graphical representation. The data demonstrate that across all 4 time points that there was no significant difference for the OR or CI for MFI as none of the quantities fell beyond the 95% CI.

Finally, for MFI at T1 the p-value was 0.066 which is above the standard alpha of 0.05, but it was below the threshold of this study of 0.10. Because of this relatively larger p-value, the OR crossed the 1.0 threshold which normally indicates a non-significant finding, however, since this was an exploratory study, the CI crossing 1.0 is still considered a significant finding.

Table 4.14 MFI odds ratio and confidence interval at all time points

	Odds ratio	**95% Confidence Interval**
Time point 1	1.10	0.99 to 1.22
Time point 2	1.14	1.03 to 1.26
Time point 3	1.12	1.00 to 1.24
Time point 4	1.15	1.02 to 1.31

Figure 4.9 Graphical representation of odds ratio and confidence interval for MFI across time.

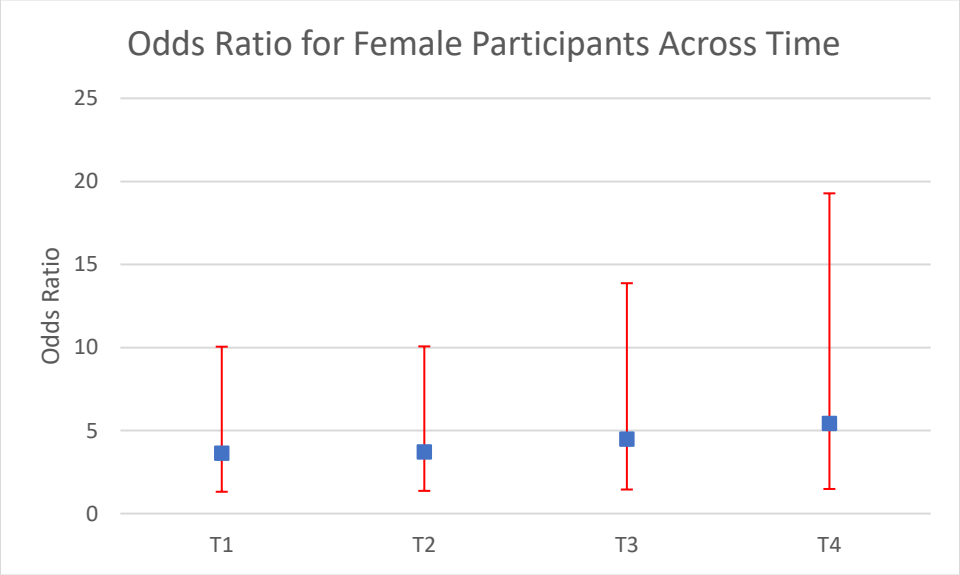


Sex was evaluated across time. Female sex as a main effect was a significant predictor of persistent symptoms group membership at all time points; see Table 4.15 for sex OR and 95% CIs at each of the four times and Figure 4.10 for its graphical representation. An evaluation of the data in the graph and table demonstrated that there was no significant change for sex in terms of OR or 95% CI across time in those with persistent symptoms.

Table 4.15 Sex odds ratio and confidence interval at all time points

	Odds ratio	95% Confidence Interval
Time point 1	3.63	1.31 to 10.05
Time point 2	3.71	1.37 to 10.07
Time point	4.48	1.45 to 13.87
Time point 4	5.42	1.48 to 19.82

Figure 4.10 Graphical representation of odds ratio and confidence interval for sex across time



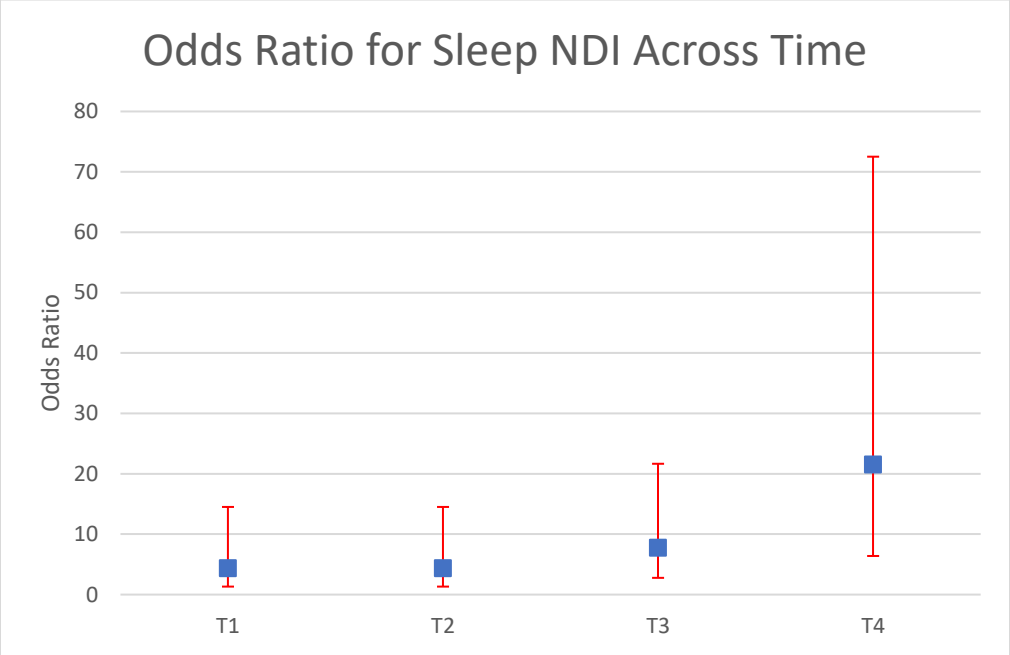
For sleep NDI, the data demonstrated it as a significant predictor of group membership at T1, T3, and T4 for its main effect. Because sleep NDI was non-significant at T2, the OR and 95% CI from T1 were carried forward as a placeholder. See Table 4.16 for OR and 95% CIs for sleep NDI, and Figure 4.11 for graphical representation.

For evaluating differences across time, there were no differences noted for the CI or the OR when compared to the immediately preceding time point, but at T4 the OR was only 0.13 away from being classified as a significantly different from T3 as the OR at T4 was 21.53 and the upper boundary of the 95% CI at T3 was 21.66. However, there was a statistically significant increase in the OR when T4 was compared to the 95% CI at T1. As this was an exploratory study, these results do lead to questions regarding the influence of sleep and persistent WAD symptoms over time that may need to be more thoroughly addressed.

Table 4.16 Sleep NDI odds ratio and confidence interval at all time points

	Odds ratio	**95% Confidence interval**
Time point 1	4.38	1.32 to 14.51
Time point 2	4.38	1.32 to 14.51
Time point	7.75	2.77 to 21.66
Time point 4	21.53	6.39 to 72.51

Figure 4.11 Graphical representation of odds ratio and confidence interval for sleep NDI across time



4.5 Summary

This chapter provided the detailed results of the statistical analysis for the study "Neck Muscle Composition in Persistent Whiplash Associated Disorder: Relationship with Neck-Related Disability." The normality tests used to assess the data were described. The overall view of the data at baseline is that the variables among the participants were normally distributed. Beyond the normality tests, there was the discussion for the use of two groups (persistent symptoms and recovered) in the analysis instead of the more commonly described three-group model. As well, the general view of the study was described as exploratory in nature; thus, to minimize Type II errors, the use of an elevated alpha level (0.10) was discussed in terms of determining significance levels for the predictor variables.

Regarding data at baseline, there were several variables that were different between the groups. These variables included MFI, NDI scores, sleep NDI responses, as well as the sex of the participants in the groups. The first research question aimed to identify which variables were significant at baseline for determining group membership at 12 months. Initially, the variables sex, MFI, and sleep NDI were found to be statistically significant for main effects, and no interaction effects were found to be significant.

The second research question evaluated which variables, identified in the primary research question, changed in magnitude or direction over time. In general, the main effects for the variables MFI and sex were consistent across the time points, without any statistical significance noted. For sleep NDI response, there was a large increase in the odds ratio at T4 that was significantly different when compared to T1, but when compared to T3 it was not statistically different. As the effect size for the sleep NDI variable increases across time, it may be a variable of interest for future studies. As well, across time, there were no interaction effects that were found to be significant at any of the time points. Finally, the pseudo R^2 for the overall model prediction of group membership increased from 0.27 at T1 to 0.56 at T4, as well as the correct prediction of both the recovered group and those with persistent symptoms.

Chapter 5: Discussion

5.1 Introduction

The purpose of this chapter is to discuss the data presented in Chapter 4, specifically, recognizing how the data fits and expands our understanding of currently published literature regarding whiplash and its associated complexities. The first portion of the chapter reports which variables were found to be significant predictors of persistent symptoms at baseline. The second part of the chapter evaluates these variables across multiple time points, individually and as a complete model, both in terms of their effect sizes and their ability to discriminate between the two groups. Finally, the clinical significance of the findings are discussed, future research lines are proposed, and the limitations and delimitations that were present in the study are addressed.

5.2 Specific Aim #1

The primary aim of this study was to establish the relationship between the independent variables (IV), body mass index (BMI), age, sex, cervical muscle fatty infiltrate (MFI), numeric pain rating scale (NPRS), and sleep response on the neck disability index (sleep NDI), and the dependent variable (DV) group membership following a motor vehicle crash (MVC) of persistent symptoms or recovered.

5.2.1 Non-Significant Variables

From the results, the individual factors that were reported as having no significant difference between the groups at baseline as well as being non-significant in terms of prognosis of final group membership were BMI, age of the participant, and the participants score on the NPRS. In the literature, there is conflicting evidence as to the influence of the factors BMI and age on outcomes of people with whiplash associated disorder (WAD) as some studies report a

predictive relationship while others claim no association between the variables.[151-153] However, for the NPRS, the results reported here contrast with multiple studies which report the variable as a strong predictor of a poorer outcome.[77,151-153]

It has been theorized that those with an increased BMI may report higher levels of persistent symptoms following an MVC due to systemic inflammation, structural or mechanical changes due to body habitus, or psychosocial variables associated with being obese.[111] As well, advancing age has been thought to be related to increased symptoms based on the potential for diminished bone health and degeneration of the structures within the neck that may become symptomatic following the trauma from the MVC.[154]

Multiple systematic reviews have evaluated BMI and age in regard to WAD and persistent symptoms and reported them to be either non-significant as a predictor of group membership or described inconclusive results.[151-153] The results from this study support the notion of both variables being non-significant predictors of group membership. Both variables were assessed at each of the 4 time points (T1 1-week; T2 2-weeks; T3 3-months; and T4 12-months) and found to be non-significant at all points except for BMI at T4 when an unstandardized beta was reported of -0.182 and a 95% confidence interval (CI) of -0.35 to -0.013. Interestingly, with the reported beta below zero, this result implies that an elevated BMI is protective against the development of persistent symptoms associated with WAD, but this goes against previously published theories which hypothesized an elevated BMI as a predictor of WAD.[152] It is unknown exactly why the beta value would be less than zero, but it is possible that a larger body mass can protect the deeper structures from trauma that may occur in the MVC. However, given that BMI was significant at only one time point, it does not appear to be a strong

result and it is possible that the result may be a statistical error related to the multiple tests run at all time points.

One final point about BMI that was of interest in this study was that its correlation with cervical MFI was significant but weak with r values between 0.19 to 0.24. It could be concluded that a person with an elevated BMI would also have an elevated level of MFI, but that was not the case for these participants. With this weak correlation, along with the result that demonstrates that the levels of MFI that were present at baseline did not significantly change over the course of the 12 months, it appears MFI may be a predetermined individual trait that may influence the DV. The MFI present at baseline may be a result of genetics, hormones, or any numerous metabolic interactions;[32] the implications of MFI on WAD will be discussed later in this chapter.

Within the inclusion and exclusion criteria for the study was the age range from 18 to 55. The theory behind the cutoff at 55 was to try to minimize the potential influence that degenerative changes may have on neck pain. It is not uncommon for people to have abnormal findings of their cervical spine on Magnetic Resonance Imaging (MRI) while being aymptomatic.[111,155] But it is unknown if these abnormal findings on MRI can influence neck pain or disability following an MVC. There have been published reports that there is no relationship between degeneration and pain,[111] whereas others claim an association between degeneration and neck pain, especially following a trauma.[154] It has also been noted that as we age, especially above 55, there is the increased risk of decreased bone density, and this may affect symptoms that occur following a trauma such as an MVC.[154]

Alternatively, it has been reported that neck pain in the general population tends to peak for people in their 30s and 40s.[156] However, if degeneration is a primary cause for pain, then it must be understood why people in their 60s and 70s, who have more advanced degeneration,

have a lower incidence of neck pain. Given that the role of degeneration and its influence on neck pain is unresolved, the cut-off of 55 years old was used in this study to mitigate any potential influence it may have. The results of this study were therefore based on the participants being between 18 to 55 years old, and the variable was found to have no significant influence on group membership at 12 months, but it is unknown if these results hold true for people over the age of 55.

Although the results from BMI and age appear to align with previously cited literature, the results related to the NPRS were quite distinct. In this study at baseline, the NPRS was found to be neither different between the groups nor significant for predicting group membership at 12 months as the variable had a reported p-value of 0.19 in the final model at baseline. This finding of non-significance as a predictor of future symptoms is in contrast to published articles and systematic reviews that report, with strong confidence, baseline NPRS scores as a strong predictor of chronic WAD symptoms ranging from 6 to 12-months post-MVC.[17,77,103]

A meta-analysis that evaluated high initial pain intensity (defined as pain rated as equal to or greater than 55 out of 100 on one scale or 6 out of 10 on shorter scales) as a predictor of future pain and disability from 11 separate studies reported an OR of 5.61 and a 95% CI from 3.74 to 8.43.[82] However, within these 11 studies, 2 reported a 95% CI that crossed the 1.0 threshold, and when this level is passed the results are often deemed non-significant. When an OR is greater than 1.0 it means that the variable is a predictor of the outcome, but when it is lower than 1.0 it indicates that if the variable is present the participant is protected from the chosen group associated with the DV.[157] Thus, if the 95% CI crosses the 1.0 threshold, it implies that the variable is both protective and predictive. Since, the variable cannot be both, the finding is considered non-significant.[157] Of the other 9 studies in the meta-analysis, there were generally

large CIs. Of the studies that did not cross the 1.0 threshold, the tightest 95% CI was 6.32 to 12.38, and the widest was 2.54 to 30.72.[82] With larger CIs, it is difficult to establish the precision of the reported OR based on these studies.

Because of the differences in the results from this study when compared to previously cited literature it is important to understand why they exist; in this case, it may be due to the way the groups were allocated. Most studies will use a 3-group model for categorizing participants following an MVC (recovered, mild, and moderate-severe)[19,77] and the NPRS is generally associated with the moderate to severe group. Studies that use a 2-group model have a score on the Neck Disability Index (NDI) of 30% or above as one group (moderate to severe) and below 30% for NDI scores as the other group (recovered to mild).[20] In this study, only 2-groups were used, but the cutoff of the NDI was 10% for those with persistent symptoms. Therefore, it is plausible that had the NDI of 30% been used in this study as the criteria for the DV, then the NPRS may have been found to be a strong predictive variable.

Another possibility for the differences between this study and some of the previous literature may be accounted for based on the outcome variable. In this study the NDI, a disability scale, was used as the DV, whereas in the previously cited meta-analysis, only 3 of the 11 studies used a disability scale as the DV: the other 8 used pain as the DV. It should be considered whether it is appropriate to use an IV at baseline to describe the same variable later but at that time calling it the DV, but this is what occurred when pain was used at baseline to predict pain at the end of the study in those previous studies.

Finally, of the 3 studies that used the disability scale as the DV, all 3 had an OR for the NPRS higher than the meta-analysis average and the 95% CIs were quite wide, and one of the studies had the CI crossing the 1.0 threshold. With a wide CI, it is unknown how precise the

findings are in relation to the DV and may lead to a larger question about the relationship between pain and disability.

According to the International Association for the Study of Pain, *pain* is defined as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage or described in terms of such damage."[158] A person with a disability, as defined by the Americans with Disability Act, is someone who has "a physical or mental impairment that significantly impairs one or more major life activities."[159] It is not uncommon for someone who has pain to also be limited in any or all of their daily life activities, as the activities can consist of caring for themselves or family members, working on their computer, sleeping, or participating in recreational activities. The difficulty in assessing disability amongst a population has to do with the notion that people have different activities that they deem important. Thus, the limitations caused by the MVC can have a greater or lesser influence on their life and potentially their personal definition of disability.

If the NDI is used as a measure of disability, we must assume that the items assessed are part of that person's daily activities as eight of the ten questions have to do with these; the other two questions are about headaches and their pain. The NDI, in its attempt to evaluate disability, asks broad questions regarding these activities. For example, the NDI tries to quantify how much their sleep is disturbed by asking specifically about their quantity of sleep loss, but it does not asses the overall the quality of sleep. As well, the NDI evaluates a person's headache frequency and intensity, but it neglects to quantify the duration of the symptoms; or a question about the person's ability to perform their daily work and the quantity performed without any assessment of the difficulty performing the tasks. Although the NDI attempts to quantify disability, as a tool

it may be too blunt, as there are many reasons why people may be having difficulty with their daily activities.

In previous studies using the NDI as the DV, often the score of significance was an NDI over 30%, and at this range it is likely that there is at least a modest amount of pain, and hence possibly the reason that the NPRS was a predictor for those with WAD. However, in what is normally considered the mild group, NDI 10 to 28%, there may be some disability such as difficulty with reading, driving, or concentrating, but there may not be much pain because WAD can present as a constellation of symptoms and not all are pain based. Symptoms associated with WAD have been stated to consist of not only neck pain, loss of range of motion, and tenderness to palpation, but also dizziness, cognitive difficulties, unsteadiness, depression, and anxiety, to name a few.[2,3] All these symptoms have the potential to influence someone's daily activities without any pain. Thus, the connection between neck pain and disability in those with WAD may not be straightforward and may be a strong reason for the inability of the NPRS to be used as a predictor of disability following an MVC in this study. As a result, when bringing these two groups together, as was done in this study with persistent symptoms greater than 10% on the NDI, the predictive capability of the NPRS may be lost at baseline and across the subsequent time points.

Using the NPRS as a prognostic variable to determine pain intensity may be appropriate as several systematic reviews and a meta-analysis demonstrate its effectiveness when pain is the DV. But using the NPRS as a predictor of future disability may need to be reexamined based on the limited results in this and other studies regarding the questionable relationship between the two variables.

Prior to examining the rest of the variables and their relation to the current literature, it would be beneficial to discuss some of the limitations in trying to describe the context of the results with published data. The primary difficulty in comparing these results has to do with the heterogeneity of the collection and reporting of previously published data.[151-153] For example, studies talk about older age but do not often quantify what is meant by the term *older*.[151] Also, the reporting of variables are sometimes referred to in the text as significant or non-significant while providing only *p*-values and not presenting any effect sizes; hence, it is unknown if the results have small, medium, or large effects.[152]

The other dilemma in comparing data is in the use of the term *recovery*. Across the literature, there are multiple ways to report recovery which include, but are not limited to self-reporting of pain, headaches, limited strength or range of motion, levels of fatigue or confusion, or scores on various outcome measures such as the NDI.[160] The variable used as the determination of recovery has the potential to influence the statistical analysis that the authors used in their study, which makes comparisons between studies that have different outcome variables difficult.[160] In this study, an attempt has been made to evaluate outcomes related to the NDI as it is the most widely used outcome measure in clinics and research studies as well as having adequate psychometric properties for this population.

5.2.2 Muscle Fatty Infiltrate

At baseline, the mean MFI was statistically different between the two groups, but this is in contrast to previous reports which showed at 1-week and 2-weeks post-MVC that MFI was not different between the groups based on their long-term prognosis.[19,20] However, there were differences between this study and the previous results in terms of both measurements and comparisons of groups that may explain the contrasting results.

In a report in 2011, MFI as measured by MRI demonstrated no difference between the groups of people with WAD at 2 weeks. But by 3 months, there was a significant difference between the groups that developed moderate-severe symptoms and those that recovered, and by 6-months post-MVC these differences became even larger.[19] However, there were several distinct differences between that study and the results reported here that may explain some of disparity.

In the previous study, they used the more traditional categorizing of participants into 3 groups based on their NDI score: 0 to 8% recovered, 10-28% mild, and greater than 30% for moderate-severe, and in that study the greatest difference for MFI was with the moderate-severe group when compared to the other two groups. The MFI for the moderate-severe group was significantly elevated when compared to both the mild and recovered groups at both 3 and 6 months despite all the groups having equal quantities at baseline. In this study, there were only 2 groups: recovered 0-8% and persistent symptoms greater than 10%; however, when participants were categorized into the 3 groups previously described both the mild and severe groups were equivalent but were significantly different from the recovered group at baseline. Therefore, it can be theorized that group categorizing can explain some of the discrepancy between these results.

The second difference between the studies had to do with the collection of data via MRI. In the previous study, there was no mention of the strength of magnet used in the machine 1.5T vs 3.0T, and the researchers reported using a T1 image for quantification of MFI. In this study, a 3T MRI was used and the 3D multi-echo Dixon fat/water imaging method was used to collect data on MFI in the participants. A comparison between the T1 and Dixon method showed that they are equivalent in their quantification of MFI,[127] but the Dixon method results in faster acquisition times of the tissue which results in in less chance for artifacts to be displayed.[10] With

improved imaging acquisition, it is theorized that using the Dixon method to visualize MFI may allow for differences between the groups to be viewed earlier when compared to the standard methods using T1-weighted sequences.[20] It is this image acquisition that may explain why differences were viewed at 1 week in this study whereas they were not visible at 4 weeks in previous studies.

In the second study in 2015, researchers reported the use of the dual-echo fat/water Dixon method to generate the images that allowed for the calculation of the MFI, but again, there were other differences which may have led to the contrast in results of MFI at baseline. First, in the 2015 study, there was a 60% rate of non-compliance related to image acquisition with 20/89 participants stating that they could not commit to the 3-month timeline and 26/89 who consented but did not follow up for their images. It is unknown if these dropouts affected the final statistical analysis, but given that some authors state that non-compliance greater than 20% can lead to bias, the 60% in this study does have some cause for concern.[161]

The other primary area of discrepancy has to do with the categorization of the participants. In the 2015 study, participants were categorized into 2-groups based on their NDI scores at 3 months, but these two groups were recovered-mild symptoms with NDI of 0-28%, and moderate-severe symptoms with NDI greater than 30%. Categorizing the participants in this manner is in almost direct opposition to the way it was done in this study. In the standard 3-group model which classified mild symptoms as 10-28%, the 2015 study combined the mild group with recovered group, but in this study the mild group was joined with the moderate-severe group. The reasoning for combing the groups as they did, recovered and mild, was not specifically listed, but in the current study, the grouping of mild and severe was based on the similarities between the two groups and their distinct difference from the recovered group. As

well, it is theorized clinically that if patients have an NDI of 20%, they could easily be seeking treatment for their limitations and they may not consider themselves recovered. Thus, in this studies' 2-group model, we were estimating when someone might begin to seek treatment based on their NDI scores and then place all those people in a single category. Finally, in order to minimize a Type II error (false negative) and to adjust for the large standard error (SE) that has been reported for the NDI, it was deemed more appropriate to combine the mild and severe rather than the mild and recovered.

In this current data set, the mild and severe groups showed MFI levels that were similar and did not have any statistical difference between them at baseline or at any of the follow-up time points, but both groups were statistically different from the recovered group at all 4 time points. If the analysis had been run identically to the one in 2015 with the previously described categorization, there would have been no statistical difference between the groups.

One final difference between the current study and these two previous studies has to do with their interactions and analysis of the mediation of post-traumatic stress disorder (PTSD) and its potential influence on WAD. In this study, PTSD was not part of the analysis; thus, it is unknown how much of an effect it may have on the quantity of MFI, any potential interactions, or the long-term outcome for those with WAD. PTSD has been reported to be a mediating factor between MFI and chronic WAD,[20] and in this study it may explain some of the differences between the groups as well as being a pre-existing vulnerability for those that develop chronic symptoms. It is unknown how the presence of PTSD may influence the levels of MFI, but a few theories put forth include an increased sympathetic response that can alter muscle tissue health or the oxidative stress can create fibrotic tissue and limit muscle contractility.[19] These changes in

muscle morphology can potentially influence the biomechanics of the cervical spine and perpetuate persistent symptoms.

It is not understood exactly how MFI will influence the biomechanics at the joints, but several theories exist. Similar to visceral fat, MFI is believed to release proinflammatory cytokines following a trauma, such as an MVC, that have the potential to influence muscle contractility.[32,19] As well, MFI is also believed to alter the muscle fiber orientation and decrease the overall extensibility of the tissue which can explain decreases in strength and force production for these muscles.[32] As MFI is deposited in the tissue, a form of pseudohypertrophy is seen within the muscle that is not visualized in healthy controls.[162] The MFI is found more in the deep stabilizing muscles such as the longus coli in the flexor musculature and the multifidi and semispinalis cervicis in the posterior cervical muscles.[18,34] It is possible that any or all of these factors can influence the joint mechanics and motor control of the cervical spine which may lead to persistent symptoms.

The mechanical deficits of the cervical musculature following an MVC that have been reported in the literature include decreased endurance and activation of the deep cervical neck flexors, altered cervical muscle activation with upper extremity movements, increased joint position errors, and decreased range of motion.[163-165] However, these deficits are also present in people with insidious onset neck pain[166] despite those with WAD having elevated levels of MFI when compared to the insidious onset group.[12] But, it is believed that those with WAD have motor deficits that are greater in magnitude[165,166] and these differences may explain some of the symptoms that present in this group.

Given that both those with insidious onset chronic neck pain and with WAD demonstrate altered movement patterns, it is theorized that pain has an underlying influence on motor control

within the cervical spine.[166] However, in addition to the pain, MFI may also influence muscle activity and further propagate some of the disordered movement patterns that are present in those with WAD. The combination of both pain and elevated levels of MFI may explain why those with WAD have greater changes associated with their movement patterns when compared to healthy controls and insidious onset neck pain.

In the anatomy of the cervical spine, there are numerous layers of muscles that contribute to the mobility and stability of the structure.[167,168] As well, there is also significant overlap in the musculature related to motor control that allows for the precision of movement.[169] Although MFI can likely affect the motor control and precision of movements, it is not expected to be the singular factor that will determine if someone has persistent symptoms associated with WAD. But, unlike some other variables associated with WAD symptoms, MFI is a variable that can be altered with treatment and specific exercises.[96] Thus, it is an avenue for evaluation and treatment that has the potential to improve outcomes for those that are diagnosed with WAD.

In this study, at baseline, MFI was reported to have an unstandardized beta value of 0.10 and a 95% CI of 0.00 to 0.20. What this means is that for every 1% increase in MFI, the log odds of a person developing persistent symptoms will increase by 0.10. At baseline, those that ended up in the recovered group had an average quantity of MFI of 17.8% while those in the persistent group had 20.95%, a difference of 3.15%. Therefore, if each 1% increase in MFI generates a 0.10 increase in the logs odds of a person developing persistent symptoms with an average difference of 3.15%, a new beta value can be calculated as 0.10 * 3.15 = 0.315. With this unstandardized beta value, an OR of 1.37 is reported for a participant that has MFI that is 3.15% higher than the recovered group. Thus, this person will have approximately 1.37 times the odds

of developing persistent symptoms when compared to the recovered group based on their quantity of MFI.

5.2.3 Females

At baseline, in this study, the female-to-male ratio was reported as being statistically different between the groups of persistent symptoms and recovered at 12-months post-MVC. In the recovered group, there was an N of 37 participants with 21 females and 16 males, while in the persistent symptoms there was an N of 60 with 51 females and 9 males. Beyond the differences at baseline, the female sex was found to be a statistically significant predictor of persistent symptoms at 12 months with an OR of 3.63 and a 95% CI 1.31 to 10.05 These results are not uncommon in the literature, but they are also not definitive, as multiple systematic reviews have looked at the sex of the vehicle occupants and found results varying from inconclusive, to no effect, to a significant moderate effect of the variable.[108,151,152]

There are multiple theories as to why the variable does not provide consistent results, with one being that it is simply not a strong variable of influence; accordingly, the results will fluctuate from study to study. In one meta-analysis, based on 14 previous studies the calculated OR was 1.64 with a 95% CI of 1.27 to 2.12,[151] which means that for every one male that develops persistent symptoms, 1.64 females will develop them (to put it in a larger scale, for every 10 males who develop symptoms approximately 16 female will as well). These reports, although significant, are not large enough to provide clinical utility. With the OR not being large, it is easy to understand how some studies can report significant effects while others cite no relationship.

A second option for explaining why females represent a greater percentage of those with persistent symptoms has to do with the use of health care services for both males and females. In

general, with all other variables accounted for, females will seek and receive more health care services than males.[170] It is unknown exactly why females demonstrate an increased use of health care services when compared to males. One theory is that they are more aware of their symptoms and are more interested in maintaining their health, whereas others report that females experience more health-related issues (WAD could be in that category) that require medical attention and treatment.[170]

The opposite side of the argument also needs to be examined, as rather than seeing women as having an increased use of health services, it is just as possible that men underutilize them.[171] There are no definitive reasons as to why a man may not seek help for his health care needs, but some presumptions are that men view seeking help as an unacceptable option; they view themselves as tough and not needing care, or they see that seeking help exposes their vulnerability and they wish to keep that part of themselves hidden.[171] It is certainly possible that it is a mix of both the females and males use of health care services, or lack thereof, that may influence the numbers of females that report with WAD. With all of this in mind, in this and many other studies, as participants are recruited from hospital emergency departments (ED), if a group of people are less likely to go to the ED (males) then they are less likely to be recruited and studied and may be underrepresented in the final analysis.

A final possibility for why there are more females in the persistent group has do with multiple factors that are associated with their anatomy when compared to males such as decreased muscle or body mass, or decreased height, both of which theoretically could influence outcomes.[112,172] Within these theories is the notion that seats and safety features are generally developed based on the size of the average male driver regarding the stiffness of the seat or the height of the head restraint.[112,172] These factors could alter the force transmission of the vehicle

to the occupant during an MVC and result in greater rates of injury and symptoms for those who do not meet the size standards for designing the cars. However, a systematic review reported high confidence that seat position had no influence on outcomes, but it was noted that the seat positions were self-reported and not objectively measured.[17] On the other hand, in a study that objectively assessed head-restraint position, the authors reported that 93% of automobiles had head restraints that were less than optimally positioned,[173] thus, the self-report method may not be appropriate when assessing this variable.

Finally, some females may be the same size, or larger than males, and may have fewer injuries than their smaller female counterparts which could result in fewer injuries for this subgroup. Given that the previously mentioned systematic review did not report sex differences in relation to seat position, it is possible that the size differential between larger and smaller females, and the potential for more injuries in the smaller group, may be part of the reason why the role of female sex and its influence on persistent symptoms is difficult to fully ascertain.

With all these factors in consideration, small effect size, differences in health care utilization, and differences in anatomy, in this study, at baseline the variable sex had an OR of 3.63 with a 95% CI of 1.31 to 10.05. The OR of 3.63 is twice as large as the OR reported in the previously cited meta-analysis of 1.64. Some of the difference between these OR may have to do with how persistent symptoms were reported in this study; NDI over 10% was considered persistent whereas other studies often used the 30% threshold. With a lower threshold for NDI, more people were classified as having persistent symptoms, 61%; if the 30% threshold were used, only 17% would have reported persistent symptoms at 12 months. Given that females are more likely to use health care services and that the participants were recruited from an ED, there will likely be more females in the study, and it can be understood why the OR for the female sex

may be higher in this data set. In this study, if the 30% threshold were used than 16 participants would be classified as having persistent symptoms and the male-to-female ratio would be 1 to 4, whereas using the 10% NDI threshold the ratio was 1 to 5.6; using the lower NDI threshold likely explains some of the increased reported OR in this study.

The purpose of the OR is to estimate the size of the effect the IV may have on the DV, whereas the CI is a report on the precision of the OR.[157] In the results, looking at the effects of female sex on group membership, the 95% CI ranges from 1.31 to 10.05, which is not a precise measurement for the OR. The CI is stating that, with 95% confidence, the true OR may be as low as 1.31 (which is close to the previously cited meta-analysis) or as high as 10.05. We simply do not know the true value of the OR based on the data provided. Although the OR of 3.63 is statistically significant and generally larger than previously cited reports, the precision of this estimate is less than optimal and needs to be considered when evaluating it. Be that as it may, the sex of the occupant at baseline is still considered a significant variable in predicting long-term outcomes with an effect size at least as large as the previously cited meta-analysis, if not larger.

5.2.4 Sleep Scores

For assessing sleep disturbances, the participants were categorized into two groups: one, no sleep disturbances due to the MVC, and two, any sleep disturbances due to the MVC. In this study, the sleep response was recorded from the item response on the NDI and was found to be significant in terms of differences between the groups, as well it was found to be a predictor for determining group membership at 12 months. In terms of the differences between the groups, the Mann-Whitney U-test reported a significance level of 0.027 with the recovered group reporting a median sleep response of 1 and the persistent group reporting a response of 2. Although, these numbers were statistically significant, it is unknown if they are clinically significant. For the

prediction of 12-month group membership, the sleep NDI response reported a *p*-value of 0.016 and an OR of 4.38 with a 95% CI 1.32 to 14.51, which means that if someone had any disturbance to their sleep due to the MVC, their odds of developing persistent symptoms increased 4.38 times.

In looking at the median responses and the scores of 1 and 2 for each group, clinically it is difficult to determine if there is much of a difference between them. Scoring on the NDI for sleep loss is provided in groups with 0 being no disturbance, a score of 1 is less than 1 hour of disturbed sleep, a score of 2 is between 1 and 2 hours of disturbed sleep, and a score of 3 is between 2 and 3 hours of disturbed sleep. Given the sleep NDI response for an individual, a score of 1 can span from a few minutes of sleep disturbances up to an hour, whereas a score of 2 can range from 1 to 2 hours. As there is some overlap between the categories, (the 1-hour mark being in both scoring categories) the difference in the quantity of sleep loss may be nearly 2 hours or as little as a few minutes, and as these scores are self-reported it is unknown just how much difference in sleep loss there is between the groups.

The potential lack of clinical significance may have to do with the timing of the initial evaluation. Perhaps it was too soon for the differences to be fully realized. The primary recruitment for participants in this study was from the ED within 1 week of the MVC; thus, it can be hypothesized that the MVC was significant enough to warrant an immediate evaluation with a medical doctor. With this in mind, most participants will likely have some level of stress and anxiety related to the ED visit even if they are provided a clean bill of health, and it may be worse for those that continue to feel some symptoms. Although no research could be found that looked at sleep and pain associated with an acute (less than 1-week) MVC, there was research associated with sleep and recent trauma.

In literature evaluating sleep and pain associated with trauma, with the traumas reported as recent orthopedic post-operative procedures, cardiovascular, and cancer treatments, as well as burn victims, the influence of pain and sleep disturbances were reported to be bidirectional.[115,117,174] With a bidirectional relationship, the pain associated with the trauma can influence the quality of the sleep, while at the same time the quality of the sleep can affect the pain.[174] However, it is believed that sleep loss and decreased quality of sleep influences pain more than the pain affects sleep.[117,175] Unfortunately, the effect size is not stated and it is not well understood if it is a direct relationship or if there is a mediating factor.

Although it is unknown exactly how pain and sleep interact in the acute state,[117,174,175] it is theorized that post-operatively there is an endocrine, an autonomic, and an inflammatory response with the release of various cytokines and interleukins that can interrupt both slow wave sleep (SWS) pattern as well as rapid eye movement (REM) sleep.[174] The stress response system can also release cortisol and adrenocortotropic hormones in response to the trauma and these also have the potential to influence the quality of the sleep cycle.[174]

It is possible that the initial stress and trauma from the MVC can have a hormonal and endocrine response which can influence the quality of the participant's sleep. Since it is likely that most of the participants have some level of stress associated with the MVC, the sleep disturbances between the groups may only be different by a small margin at baseline.

In evaluating the OR and trying to understand the width of the 95% CI (1.32 to 14.51), the tool used to assess sleep disturbances should be evaluated. It is possible the sleep response on the NDI may be too blunt of a tool to fully articulate the predictive capabilities of sleep disturbances on persistent symptoms. The scoring for sleep disturbances on the NDI are based on a 0 to 5-point scale with increasing quantities of lost sleep reported at each level (A score of 0 is

no trouble sleeping, a score of 1 is less than 1 hour of sleep loss, and a score of 5 is 5 to 7 hours of sleep loss). The scoring provides a general overview of quantity of sleep, but a tool such as the Pittsburgh Sleep Quality Index (PSQI) may be able to distinguish subtleties that are not apparent when only looking at quantity.[176] The PSQI asks 19 questions and evaluates seven distinct components of sleep including (1) sleep quality, (2) sleep latency, (3) sleep duration, (4) habitual sleep efficiency, (5) sleep disturbances, (6) sleep medication, and (7) daytime dysfunction.[176]

In looking at post-operative patients, regarding pain and sleep loss, researchers did not demonstrate simply decreased quantity of sleep, but they showed specific patterns with deficits in both SWS and REM.[174] REM and SWS are important in terms of both increasing the quality and restorative factors that are associated with sleeping, which is primarily maintaining homeostasis across multiple physiological systems.[51,114,115] Thus, if following an MVC people are having diminished quality of sleep by alterations in their REM or SWS patterns, this can lead to not feeling rested after waking in the morning or to daytime drowsiness, even though their quantity of sleep is not significantly changed. Therefore, a tool such as the NDI that is only looking at quantity of sleep may not be able to detect changes immediately after the MVC and a more nuanced assessment of sleep architecture, such as the PSQI, may be needed to fully elucidate the problem at the 1-week time frame.

Finally, as the scores for sleep disturbances were pulled from a question within the NDI (and the NDI score at 12 months was used as the DV to determine final group membership), there is the concern for a significant correlation between this IV and the DV that may explain some of the effect size reported in the analysis. In assessing the relationship between the sleep NDI scores at baseline and NDI scores at 12 months, there was a significant correlation with an r of 0.23, which indicates that as one score goes up so does the other one. Although the strength of

the correlation can be considered to have little to no relationship,[136] and likely does not explain the entire effect size, it should be considered as part of the factor for the result. This correlation will be considered further when looking at sleep scores at future time points.

5.2.5 Interaction Effects

The purpose of looking for interaction effects was to determine whether the degree to which significance or effect size of any of the variables of influence changed based on the presence or value of another variable. Interactions were only considered between variables that were found to be significantly different between the groups at baseline; these variables were sex, MFI, and sleep NDI. Therefore, the three potential interactions were: sex by MFI, sex by sleep NDI, and MFI by sleep NDI. In the logistic model at baseline, none of the interactions were found to be significant.

In the current published literature, variables are often evaluated in isolation and their significance values and main effects are reported. There is some discussion on the interaction of variables, primarily PTSD and its interaction with variables associated with chronic WAD,[19,103] but this information is limited. In evaluating the interactions assessed in this study, none were found to be significant, but the variables simple main effects were significant, which may provide evidence for the notion that there is some component of these individual variables that may predispose a person to persistent symptoms. For example, for females it is unknown if their predisposition to persistent symptoms has to do with hormones, genetics, connective tissue strength, muscle mass, or increased use of health care services to name a few of the possible options. However, these avenues should be explored further to fully elucidate the understanding of sex and persistent symptoms.

To further understand these variables in a clinical perspective, it was of interest to see if the output related to each variable would differ if the sexes were evaluated separately. Therefore, the data was split based on sex and the logistic regression was run with the variables MFI, sleep NDI, and an interaction of sleep NDI by MFI. In the final model for each sex, it was found that none of the variables, main effects or interactions were significant, but the distribution of the data is likely the reason for these results. The data split resulted in 72 females and 25 males, but it is unknown if this nearly 3 to 1 sex ratio was adequate to generate appropriate results. With this spilt, the distribution of the participants across the groups was likely not appropriate for a statistical analysis. For males, (Table 5.1) there were 4 observation cells and in 3 of the 4 cells there were 5 or fewer participants; for females (Table 5.2), in 2 of the 4 cells there were 6 or fewer participants. With such a small N in numerous cells, it is possible that the effect of the variables split by sex was not evident.

Table 5.1 Classification table of predicted and observed group membership for males only.

Male		
	Predicted recovered	**Predicted persistent symptoms**
Observed recovered	14	2
Observed persistent symptoms	5	4

Table 5.2 Classification table of predicted and observed group membership for females only.

Female		
	Predicted recovered	**Predicted persistent symptoms**
Observed recovered	6	15
Observed persistent symptoms	2	49

To further assess the potential influence of female sex on the other variables, and to determine if it had a mediating effect on them, a logistic model was created using only MFI, sleep NDI, and the interaction of MFI and sleep NDI; thus sex was not included as part of the analysis. Running this analysis, the output was nearly identical to the one that had those variables and sex included; the beta value for MFI without sex was 0.11 and with sex it was 0.10, for sleep NDI without sex the beta was 1.49 and with sex in the model it was 1.48. It appears that the variable sex is not a mediating factor for sleep NDI or MFI on the DV.

When sex is added into the model, there is a change in the overall classification table and an improvement in the specificity of the model. The sensitivity of the complete model with sex is 85%, without sex in the model it is 86.7%, whereas the specificity with sex in the model is 62.2% and without sex in the model it is 35.1%. Therefore, by adding sex into the model, the specificity increases by 27.1% without influencing the sensitivity. Consequently, by adding the female sex into the model there is an improvement in the prediction of the model for ruling people in to develop persistent symptoms when these variables are present. See Table 5.3 for the comparisons of the models with and without sex as a variable.

Table 5.3 Comparison of 2 models with and without sex included as a variable including beta value, overall prediction, sensitivity, and specificity.

	Beta value	Percentage correctly classified overall	Percentage correct for persistent symptoms group (sensitivity)	Percentage correct for recovered group (specificity)
Model with sex included as a variable				
MFI	0.10	76.3	85.0	62.2
Sleep NDI	1.48			
Model without sex as a variable				
MFI	0.11	67.0	86.7	35.1
Sleep NDI	1.49			

Abbreviations: MFI, muscle fatty infiltrate; sleep NDI, sleep response on Neck Disability Index.

5.2.6 Model Prediction

For overall prediction of the model, a variety of analyses were performed. The first was the reporting of the Nagelkerke R^2, which is a report of the explained variance of the DV by the factors in the model. The second analysis was performed using a ROC curve and its associated AUC, which are graphical and quantitative evaluations of the model and its ability to discriminate between the two groups. Finally, there was the assessment of the model to predict the overall grouping of the participants as well as each subgroup: persistent symptoms (sensitivity) and the recovered group (specificity).

In evaluating the model for its ability to predict persistent symptoms, the Nagelkerke R^2 was 0.27; the 0.27 reported for R^2 means that 27% of the variance of the DV can be explained by these variables. Other measures for model prediction were the receiver operating characteristic (ROC) curve and its associated area under the curve (AUC) reported at 0.68; the 0.68 means that the model has a poor ability to discriminate between the two groups. Although this model by

itself is considered to have poor discrimination of the groups, the results still can provide some benefit to patients and practitioners. While it can be argued that 73% of the variance is unknown based on this model (and this amount of variance may explain why this model is a poor predictors), of the variables that are in the model, both MFI and sleep disturbances may be amenable to treatment and can perhaps be a pathway for helping to prevent some people from developing symptoms. As well, recognizing that females have higher rates of WAD, when a female presents with sleep disturbances as well as elevated MFI levels clinicians may be more inclined to initiate treatment earlier than a male who does not have any of these factors.

In assessing the ability of the model to predict specific group designations based on the presence of the variables, the reported sensitivity was 85.0% and the specificity was 62.2%. When comparing these to the previously reported CPR, there is a considerable drop in the specificity which was 99% for the CPR, and this implies that there is an increase in the false positives related to these variables. From a clinical perspective, this result means there is an increase in the number of people who do not develop persistent symptoms, but the model predicts that they will, thus, they may end up with unnecessary treatments. However, the sensitivity in the CPR was listed at 44%, so the 85% using this model will reduce the false negatives from 56% to 15%. Therefore, using this model there would be fewer people who end up developing persistent symptoms who previously would not be detected and now could end up being referred for treatment that may help to reduce the chronicity.

5.3 Specific Aim #2

The second aim of this dissertation was to determine both if and how much the variables of interest changed over pre-determined time points: 2-weeks, 3 and 12-months post-MVC. In assessing change, both the magnitude and direction of the influence were evaluated.

5.3.1 Main Effects

The main effects for the variables of interest (sex, MFI, and sleep NDI) were found to be significant predictors of persistent symptoms at all time points except for sleep NDI at T2. As well, they demonstrated no significant change when compared to the immediately preceding time point; the changes were evaluated based on their beta values, ORs, and CIs. However, the sleep NDI at T4 was significantly different when compared to T1 for both the beta values and ORs and was nearly significant when evaluated against T3. Each variable however needs to be assessed to more fully comprehend how it responds across time.

5.3.1.a Muscle Fatty Infiltrate

MFI was found to be a significant predictor of persistent symptoms at all four time points. See Table 5.4 for beta coefficient and 95% CI. To determine if a significant change occurred across time, all variables were assessed based on the mean of the OR and the 95% CI for each time point. The mean at each time point was compared to the 95% CI of the previous time point, and if it was beyond the CI than it was deemed to be significantly different. When MFI was evaluated in this way, at no time point was the mean of the MFI outside any of the previous time points 95% CI, thus it was determined that no significant change occurred across time for MFI.

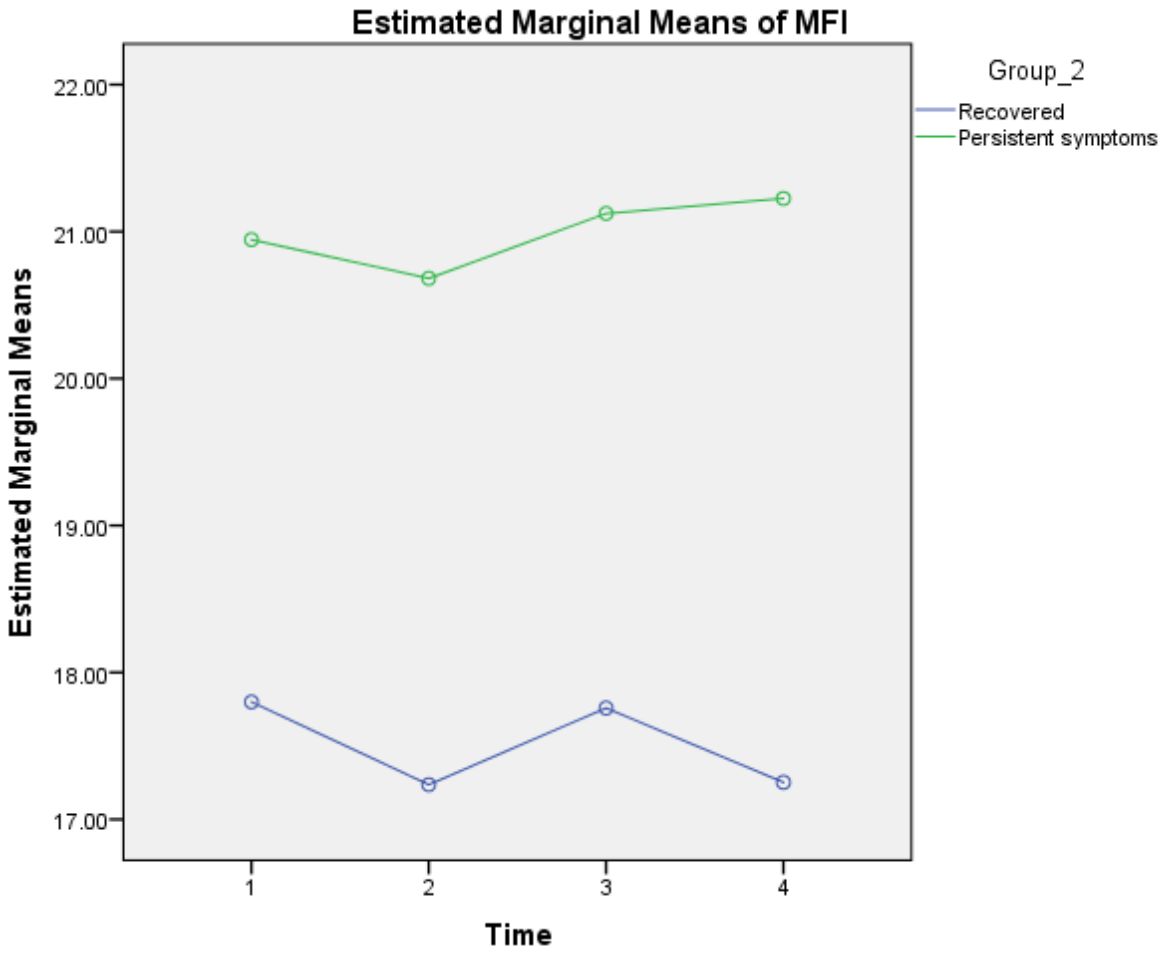
Table 5.4 MFI odds ratio and 95% CI across all time points for developing persistent symptoms

	p-value	Odds ratio for MFI	95% Confidence Interval
Time point 1	0.07	1.10	0.99 to 1.22
Time point 2	0.01	1.14	1.03 to 1.26
Time point 3	0.04	1.12	1.00 to 1.24
Time point 4	0.03	1.15	1.02 to 1.31

Abbreviations: MFI, Muscle fatty infiltrate.

To further evaluate MFI across time, a split-plot repeated measures ANOVA was used to assess if any group by time changes occurred. From this analysis, there was no statistical difference within the groups from T1 to T4 for either recovered or persistent symptoms, but there was a significant difference between the groups at all four points. As well, a group by time interaction assessing changes between the groups was non-significant. The lack of significance means that there was no difference in the way the groups changed over time when compared to each other. See Figure 5.1 for profile plot of MFI at T1 through T4 for each group.

Figure 5.1 MFI T1 through T4 separated by groups.



Abbreviations: MFI, Muscle Fatty Infiltrate.

As it was reported that there was a difference between the groups at baseline in the quantity of MFI, it can be argued that the beta at baseline is the initial influence of the variable on the DV, and it does not significantly change over time. However, this result appears to contrast with the notion that MFI can be influenced by inflammation and trauma. [12,32] The inflammatory response, although it is a tightly regulated series of steps coordinated to assist in tissue healing, does fluctuate individually based on multiple variables such as age, genetics, and tissue health.[177,178] A previous study reported differences in as little as 2 weeks in the quantity of

MFI for those with moderate to severe WAD following an MVC, and it was theorized that inflammation from the trauma may have played a part in those changes.[20] However, if trauma and inflammation were to alter the quantity of MFI, it could be expected that the overall amount of MFI would increase and there would be within-group changes. In addition, if it were to change only for one group, then it is also possible to have a between group by time interaction, where one group changed significantly compared to the other, but neither of these results appeared.

Finally, to evaluate MFI further, the quantity of MFI was assessed between the sexes. No previous studies were found that directly assessed the quantities of cervical spine MFI in females compared to males, but there were several that looked at MFI within the lumbar spine, and in most of the muscles evaluated females had a greater quantity of MFI.[179,180] In this report, females had a higher mean of MFI at 20.30% compared to 17.82% for males, but based on the size of their standard deviation, 5.9% for females and 4.6%, it appears the differences may not be statistically significant. To evaluate this variable between the sexes, an independent *t*-test was used to assess the mean differences. There was a statistically significant difference at T1, $p = 0.049$, but at the other 3 time points there was no difference between the groups. Given that the quantity of MFI was not different between the sexes at 3 of the 4 time points may explain the previous findings that showed no difference in beta coefficients whether sex was included or not in the model.

Overall, it appears that the quantity of MFI is a significant predictor of group membership at 12 months, and the greater the quantity the larger the beta coefficient and OR for that person. Although it may be a smaller when compared to other variables, it does demonstrate a consistent effect across time.

5.3.1.b Female Sex

Female sex was a statistically significant variable at all time points and demonstrated ORs and CIs that were similar at all times, with a slight increase at T3 and T4; see Table 5.5 for ORs and 95% CIs across time. Based on the mean of the OR and the 95% CI at each time point, there was no significant change across time in the effect of this variable.

Table 5.5 Odds ratio and 95% confidence interval for female sex to develop persistent symptoms across all time points.

	p-value	**Odds ratio**	**95% Confidence interval**
Time point 1	0.01	3.63	1.31 to 10.05
Time point 2	0.01	3.71	1.37 to 10.07
Time point 3	0.01	4.48	1.45 to 13.87
Time point 4	0.01	5.42	1.48 to 19.82

Female sex is a variable that has been studied in numerous reports, but the results are not consistent. In general, the effect of the variable is considered inconclusive, but the confidence in this conclusion is low.[17] However in this study, as was previously described, the DV was evaluated in a slightly different manner both in terms of using a dichotomous outcome rather than the continuous NDI score, as well as using NDI greater than 10% to represent persistent symptoms rather than the standard 30%. With these listed differences in consideration, it is possible that the OR and CI listed in this report for females is a more accurate representation of the overall prognostic power of this variable.

As the OR and CI for T1 and T2 are almost identical, it is of interest to recognize the increase in the OR and widening of the CI, especially at T4 where the upper boundary is nearly

twice as large as those at T1 and T2. The OR is simply a mathematical standardization of the unstandardized beta coefficient by taking the mathematical constant *e* and raising it to the power of the beta. Because this is an exponential function, small deviations in the beta can have a larger influence on the OR. Thus, for T4, it has a slightly larger SE for the beta at T4 (0.66) compared to T1 (0.52), but this approximately 20% increase when applied exponentially results in a larger CI at T4 when compared to T1.

However, it needs to be understood why the SE increases at T4. At T4, the groups are fully dichotomized and the full extent of the influence of the variable is known. At the previous time points, the influence of the variable can only be inferred from the data at that time point. However, at T4 because the two groups are fully separated, persistent symptoms and recovered, the two groups are at opposite ends of the continuum, which will result in a larger spread of the data and an increased SE.

From the results of this study, it appears that the sex of the participant is a significant predictor of group membership at 12 months with an increased OR for developing persistent symptoms ranging from 3.63 to 5.42. These results do not imply that men do not develop chronic WAD or that being female assures that they will develop symptoms. Instead, it needs to be recognized that in conjunction with other variables, sex is one piece of the puzzle that can help determine who will develop chronic WAD.

5.3.1.c Sleep NDI Response

The sleep NDI response is the final variable that was a significant predictor of persistent symptoms: It was significant at T1, T3, and T4 but not at T2. See Table 5.6 for ORs and 95% CIs for sleep NDI. Based on the criteria used in this study to determine a statistically significant change in the effect size, a result lying beyond the immediately preceding time points 95% CI,

there were no changes noted. However, the OR at T4 was beyond the upper boundary of the 95% CI at T1, which means that over time (from T1 and T4) there was a significant increase in the strength of the OR for being able to predict persistent symptoms.

Table 5.6 Sleep NDI odds ratio and 95% confidence interval for developing persistent symptoms across all time points.

	p-value	**Odds ratio**	**95% Confidence interval**
Time point 1	0.02	4.38	1.32 to 14.51
Time point 2	**0.14 (non-significant)**	2.30	0.77 to 6.84
Time point 3	<0.001	7.75	2.77 to 21.66
Time point 4	<0.001	21.53	6.39 to 72.51

Abbreviations: Sleep NDI, sleep response on Neck Disability Index.

The T1 scores can likely be explained by the general sleep response following an acute injury. Following an acute MVC, there will likely be any of the following: an inflammatory reaction, increased pain, stress associated with car repairs, doctors' visits, missing work, or other factors that may influence someone's sleep. Understanding why and how these factors can influence someone's sleep likely explains the reported results following an MVC. Recognizing the bidirectional relationship of sleep and chronic pain can also help understand why this variable can assist in predicting long-term persistent symptoms.

The lack of significance at T2, however, is more difficult to comprehend. To assess the variable at this time point, a McNemar test was used to determine if there was a difference in the proportions of participants between T1 and T2 who reported sleep disturbances. The McNemar test reported a non-significant result for the two groups between T1 and T2, which means that

the overall sleep NDI response was similar: At T1 82.5% (80 participants) reported sleep disturbances while at T2 it was 79.4% (77 participants). However, from T1 to T2, 6 participants transitioned from sleep disturbance to the no-sleep disturbance group while 3 participants regressed from no difficulties sleeping to having disturbances. Based on the proximity of the two evaluations, less than 1 week from the MVC for T1 and 2 weeks for T2, as well as the previously described factors that can influence sleep, this lack of significance was expected.

However, it was still not understood why sleep NDI was non-significant at this time point and thus further evaluation was required. The evaluation occurred by way of running the backward conditional logistic regression and inputting each variable individually, starting with sleep NDI. Creating a model using only the variable sleep NDI at T2 provided a final model with a significance value of 0.028 and an OR of 3.12. When sex was added into the model, the results were essentially unchanged. Yet, when MFI was added into the model the significance of sleep NDI was lost and it was removed from the model. It is unknown why these results exist or why they occur only at T2, perhaps MFI has a greater, or at least more consistent, influence on group membership when compared to sleep NDI. However, at the other 3 time points, T1, T3, and T4, the addition or removal of either of the other variables, sex or MFI, or any of the three interactions had no influence on sleep NDI as its beta value, OR, 95% CI, and significance levels were essentially unchanged. Thus, we cannot say for certain how MFI may influence the significance of sleep NDI.

Nevertheless, when the sleep NDI variable was evaluated based on other groups, female/male or high and low-MFI rates, a trend arose. The data were separated into male and female and the McNemar test was run between each time point: T1 and T2, T2 and T3, and T3 and T4. Only between T2 and T3 and only for the females was there a significant change in the

sleep NDI response rate. As well, previous research has found that MFI levels at 20.5% or higher to be a significant predictor of persistent symptoms at 12 months.[20] Thus when the sleep NDI was distributed into these groups (above and below 20.5%) and the McNemar test was run, there was a significant change in the sleep NDI response between T2 and T3, but only for those with MFI above 20.5%. Hence, in both males and those with MFI below 20.5%, there was no change in their sleep NDI across time points; this non-significance exists between T1 and T2, T2 and T3, T3 and T4, as well as between T1 and T4.

With each group the significance change occurred between T2 and T3, and it only in the groups that were previously cited as being predictors of persistent symptoms, female and elevated MFI, because of this, it appears that there is some degree of relationship. However, as these variables were worked into each model individually, it does not appear that any single variable acts as an effect mediator, nor was any interaction effect noted. Therefore, it is unknown if higher levels of MFI cause sleep disturbances or vice versa. As well, females generally have increased rates of MFI compared to males which may influence these results, so at this time we cannot determine a cause-and-effect relationship between these variables, but it is an area that could be studied in greater depth.

The final aspect of sleep NDI to acknowledge was the increasing OR and beta value reported across the four time points with T4 having the largest effect, as well it was outside of the upper boundary for the 95% CI at T1. This result implies that between T1 and T4 there was a significant increase in the OR in predicting persistent symptoms. By T4, many of the previously listed factors such as inflammation, stress, and pain have been at least reduced if not eliminated for some of the people, but others will continue to be afflicted by some of these and may continue to have sleep disturbances. Unfortunately, some of the strength of the OR for the sleep

NDI may be attributed to its being an item response of the NDI which is used to categorize the patients at T4 (this potential limitation will be discussed in section 5.6).

At T1, 80 participants were reported as having difficulties with sleeping and at T4 that number was 60; at T4 60 participants were determined to have persistent symptoms. Using a Spearman's Rho correlation between sleep NDI response and final group membership, at T1 the *r* value was 0.31 and at T4 it was 0.61, and both had significance values of 0.002 or below. Thus, it appears that as time passes, the correlation between these two variables grows in strength, which may partially explain the increased OR for this variable as a predictor. However, the OR for sleep NDI does have a large 95% CI (6.39 to 72.51); thus, it is unknown how precise this measurement is.

5.3.2 Interaction Effects

For being a predictor of persistent symptoms, there were no significant interaction effects noted at T1, or any of the other time points. Although there were no significant interactions discovered, there were a few findings noted between the variables that were of interest.

The first finding to be discussed occurred at T2 regarding how the sleep NDI variable was removed from the final model when MFI was added as a variable. It was of interest because this adjustment to the variable's significance occurred only at this time point, and at no other time did the addition or removal of either of these variables influence the final model. It would be of interest to obtain a greater understanding of sleep disturbances through a more nuanced assessment, such as the PSQI, rather than simply asking about quantity lost. For example, a relationship between sleep latency or medication usage, to name a few, may exist whereas the quantity of sleep lost may not be specific enough to fully grasp the connection between the two variables.

The second finding of interest had to do with the effect sizes of the variables. There were consistent results across the time points, especially for sex and MFI, and these results were not influenced by the inclusion or exclusion of any of the other variables. This consistency implies that the main effects of these variables are independent of the other variables, and there are no mediation effects between any of these items. Knowing that these variables have a consistent result and are independent of other variables can provide clinical utility for health care practitioners who work with this population.

5.3.3 Model Prediction

For the Nagelkerke R^2, the explained variance increased across the time points. See Table 5.7 for the reported R^2 across time. It is noted that there is a slight decrease in the statistic at T2, but this is likely due to the elimination of the sleep NDI variable, when this was forced into the model, in addition to sex and MFI, the R^2 was 0.24. As there are no CIs or SE listed for this statistic, it is impossible to determine if there is a statistical difference between the time points.

Table 5.7 Nagelkerke R^2 across all time points.

	Time point 1	**Time point 2**	**Time point 3**	**Time point 4**
Nagelkerke R^2	0.27	0.22	0.40	0.56

In an analysis of the results, the slight difference between T1 and T2 would be expected based on the short time interval between the two, yet a gradual increase across time implies that the strength of these variables' effect on the DV becomes larger. As the effect sizes of both sex and MFI remained stable across time, the increase in R^2 is likely related to the increasing effect size associated with sleep NDI. This finding is in alignment with the previously described

association between chronic pain and sleep disturbances and could be of benefit as we look for ways to treat and care for these patients. Unfortunately, as was previously described, the relationship between sleep and chronic pain is bidirectional; it is unknown at this time which way treatment should be focused to improve outcomes.

The overall percentage of those correctly predicted as well as those predicted for group membership followed a similar pattern to the R^2 in which the numbers were relatively stable across time, with a slight increase by T4, and a slight dip at T2. See Table 5.8 for overall prediction, sensitivity, and specificity across time points.

Table 5.8 Percentage of correct prediction for overall group membership, sensitivity, and specificity.

	Time point 1	**Time point 2**	**Time point 3**	**Time point 4**
Overall percentage correctly predicted for both groups	76.3	70.1	79.4	85.6
Sensitivity (percentage predicted for persistent symptoms)	85.0	86.7	88.3	86.7
Specificity (percentage predicted for recovered group)	62.2	43.2	64.9	83.8

If all participants are classified as having persistent symptoms, the percentage of correct prediction would be 61.9% (this is called the constant model). By using the variables sex, MFI, and sleep NDI, that number is increased by nearly 15% (76.3%), or it can be recognized that 38%% of those who were not correctly classified were placed into their proper group. By increasing the likelihood of correctly predicting patients, we can help determine who will

develop persistent symptoms sooner rather than later, and hopefully get them into a treatment program that will benefit them.

Based on the effect sizes of MFI and sex, it seems that the influence of these variables is similar across time in terms of predicting outcomes. In looking at T4, generating logistic models while excluding one of the variables, MFI or sex, the overall prediction, sensitivity, and specificity are changed only slightly. However, when sleep NDI is removed from the model while keeping MFI and sex, the sensitivity is unchanged, while the specificity drops from 83.8% to 40.5%, and the overall prediction drops from 85.6% to 69.1%. Therefore, it appears that MFI and sex will increase the overall prediction rate from the constant model of 61.9% to 69.1%, (an increase of 7.2%), while sleep NDI increases it from 69.1% to 85.6% (an increase of 16.5%). These results provide further evidence of the growing influence of sleep disturbances on persistent symptoms.

According to the T1 results, sex and sleep NDI have similar effect sizes with an estimated OR of 3.63 and 4.68, respectively. When each one is individually removed from the model at this time point, their influence is similar: there is a nearly 25% drop in specificity and a 7% to 9% drop in the overall prediction. At T1 with removal of MFI from the model, there is a decrease in sensitivity of nearly 7% and a 4% drop in overall prediction. These results suggest that the female sex has a greater influence on persistent symptoms earlier in the process, but by later time points, the effect is overshadowed by the growing effect size of the sleep NDI variable. Additionally, MFI, although not as large as the other variables, does demonstrate a significant and consistent effect size across all time points, which does not completely explain why some people develop persistent symptoms, but it may be a contributor to this classification.

Finally, the model's predictive capabilities were evaluated using ROC curve and AUC. The AUC for each time point, along with its associated 95% CI are listed in Table 5.9. An analysis of the AUC for each time point reveals minimal change across time in the ability of these variables to discriminate between groups as the overall results range from 0.68 to 0.73. With these results, the overall ability of these variables to discriminate between the groups is considered poor to acceptable. At no time do any of the results move beyond one of the boundaries of a previous 95% CI, therefore, there does not appear to be a significance change across time.

Table 5.9 The area under the curve and 95% confidence intervals for each time point.

	Time point 1	**Time point 2**	**Time point 3**	**Time point 4**
Area under the curve	0.68	0.69	0.70	0.73
95% confidence interval	0.57 to 0.79	0.58 to 0.80	0.59 to 0.81	0.62 to 0.83

Although the AUC for these variables is considered only poor to acceptable, it is also recognized that other variables that were not evaluated in this study, such as PTSD[103], have a recognized relationship with the development of persistent symptoms. It would be of interest to see how some of the variables from this study, especially with the adjustment in the way persistent symptoms were categorized, may interact or be affected by the inclusion of PTSD in this model. It is recognized that sleep disturbances are a core finding in people with PTSD[181], so it would be of interest to see how these two variables may interact or if the effect sizes are mediated by each other or another variable.

Because WAD is a complex condition with numerous signs and symptoms, it is not expected that a few single variables will be able to provide excellent discrimination between people who recover and those who develop persistent symptoms. However, with the results from this study, it can be recognized that these variables do have an influence that explains a portion of the overall variance of the DV, even if the discrimination between the groups is limited. With this knowledge, these variables can be included in future studies to help build and understand models that will not only help us to determine who will develop persistent symptoms earlier but will hopefully lead to improved treatment strategies for those people.

5.4 Clinical Significance

While the use of only these variables to help discriminate between groups may be limited, the information discovered in this study has clinical utility. Primarily, two of the three variables, sex and sleep disturbances, can be assessed very rapidly and consistently. For MFI, until more clinically friendly methods of assessment are available (or MFI is reported in radiological repots), and in order to minimize Type II errors, clinically it can be safe to assume that a person will require exercises to minimize MFI.

Previous studies have attempted to create clinical prediction rules (CPR) that can be used by practitioners to help determine who will develop persistent symptoms.[77] However, in the derivation of the CPR some of the factors discovered were high ratings on the NDI as well as elevated scores on the hyperarousal scale of the post-traumatic distress scale. The issue with using the NDI as a predictor variable is that it was used as the determining factor for group membership of the DV; the researchers used the baseline NDI to predict future NDI scores. In addition, hyperarousal is a subjective measure that has personal and cultural differences that may influence the CPR. For clinicians treating these patients neither variable is easily amenable to

treatment, whereas MFI quantities can be influenced by specific exercises and sleep hygiene has the potential to be improved through education and postural/mechanical changes. Therefore, these variables may provide more clinical usefulness to practitioners.

Finally, adjusting the way the persistent group was recognized placed an emphasis on minimizing the false negatives. The goal was to avoid missing people who may benefit from treatment following an MVC. As the sensitivity was never below 85%, with the use of these factors, there is only a 15% chance that people who will develop persistent symptoms will be misdiagnosed.

5.5 Future Research

The results from this study have generated an improved understanding of the effect of MFI, sex, and sleep disturbances on persistent WAD symptoms. However, there is still a significant amount of research that needs to be completed to more fully understand these variables and their influence on WAD.

Although elevated levels of MFI demonstrate a significant and consistent effect across time on persistent symptoms, it is unknown how or why some people demonstrate increased levels. It needs to be understood if MFI is a result of environmental influences such as exercise history, diet, or stress levels, if it is simply a matter of genetics, or a combination of both. By more fully understanding the origin of MFI, especially in relation to WAD, we can have a greater understanding of future treatment approaches.

MFI is not generally reported by radiologists when they read an MRI, and this is due to two primary reasons. The first is that MFI has generally not been understood regarding its influence on a diagnosis such as WAD, and without understanding its effect, it provides little to

no clinical utility. The other issue with MFI is that up until recently it was very time consuming to hand calculate the quantities present in a person's neck (the researcher in charge of calculating MFI in this study reported it took approximately 20 minutes per scan). However, recent research has used artificial intelligence for segmenting and calculating MFI within the cervical spine, and it has been reported to be as accurate as humans and it can perform the task in less than 60 seconds.[182] Therefore, as the speed of MFI calculations and the its effect on persistent symptoms both improve, it should be a variable that is reported by radiologist regularly and used by clinicians to help drive decisions.

It is recognized that not all people will have an MRI of their neck following a whiplash injury, thus having alternative ways to evaluate MFI can be beneficial to the clinician and patient. One possibility is to evaluate MFI with an ultrasound which can be used by clinicians when they have been trained and have the proper equipment. Although no studies have been completed looking at the cervical spine, one study has looked at evaluating MFI in shoulder muscles with an MRI and ultrasound, and a strong correlation ($r=0.90$) between the two was reported.[183] If a clinician does not have access to MRI results or an ultrasound, as MFI has been associated with decreased strength and motor control in the cervical spine, it would be of interest to see if clinical tests (deep neck flexor strength or joint position error) will be able to determine elevated quantities of MFI that can increase the likelihood of developing persistent symptoms. As a result, assessing MFI with ultrasound or manual muscle or joint position tests could provide benefit for patients diagnosed with WAD.

Reported in this dissertation, at baseline the variables of MFI, sex, and sleep disturbances accounted for 27% of the explained variance of the final group membership, these results were both significant, but there was still 73% that was unexplained. Therefore, a future study can look

to improve the model by seeing how the addition of known BPS variables (PTSD, anxiety, depression) will help to explain the overall variance of the final model.

Finally, from a treatment perspective, as both MFI and sleep disturbances can potentially be influenced by specific treatments, it would be of interest to see if treatments focused towards them will influence outcomes. Training specific muscles can help to decrease the quantity of MFI, but it is unknown how this influences outcomes in those with WAD. As well, sleep disturbances can be addressed from a biomechanical perspective with an understanding of sleep posture, but also sleep hygiene education such as decreasing stress and consistent bedtimes can influence both quantity and the quality of a person's sleep. However, it is not known how these will specifically influence people with WAD, thus, a future study evaluating these treatment strategies can illuminate their potential benefit.

5.6 Limitations and Delimitations

A primary limitation in this study was assessing potential participants for the presence of WAD as there is no gold standard used to make this distinction. As this a condition that is primarily based upon a person's symptoms and experiences, we are reliant on participants being truthful about their condition, their current symptoms, and their level of disability.

Beyond making a firm diagnosis, using the NDI as a tool to determine persistent symptoms was another limitation. The problem with this tool is that there is no formal definition of what it means to be recovered; as a result, it is unknown whether the NDI truly can evaluate those who are recovered compared with those who have persistent symptoms.[160] The term *recovery* has been associated with pain, headaches, fatigue, loss of range of motion, or symptoms associated with PTSD to name a few.[160] From this lack of a formal definition of recovery, there

is no standardization for what the most appropriate DV is to use or even for what levels within the DVs to use for categorizing people into a recovered or symptomatic group. For example, following an MVC, a person may have no pain in their neck or head but still feel that they have difficulty concentrating. Some practitioners may consider this person recovered, whereas others may not; it appears that determining whether someone is recovered is a matter of an individual's definition or opinion. Although the NDI is not a perfect tool, it is commonly used by practitioners and it does attempt to evaluate numerous areas of "function" to quantify a person's impairment; hence, it was deemed appropriate to use for this study.

Another potential limitation within the NDI that has been previously discussed is the way it assesses each individual subsection. For example, in evaluating headaches, it asks about frequency and severity, and rather than assessing each component individually, the two factors are placed together. One selection is "I have severe headaches which come frequently." Such a selection does not allow for the possibility of having severe headaches that come on infrequently; there is no category that allows for a mixture of the two variables. Furthermore, for the greatest impairment, the selection is "I have headaches all the time." It is unknown whether it is worse for someone to have a slight headache all day or a severe headache several times a day. However, according to the scoring schema of the NDI, the indication of headaches all the time (regardless of intensity) is classified as the worst.

In this study, the primary item response of the NDI that was evaluated dealt with sleep disturbance. It has already been discussed that this variable only assesses the quantity of sleep loss and does not evaluate the quality of sleep or the restfulness that a person obtains from sleeping. Although such a distinction may be of interest in future studies, in particular evaluating the quality of sleep following an MVC and its influence on other variables, in this study, it was

deemed appropriate to use the single question to determine sleep disturbances associated with the quantity of sleep lost.

Other limitations within this study exist such as people not participating in the study due to being unaware of it (they were not approached in the hospital or they did not see the social media posts), they lacked the transportation to travel back and forth for the serial imaging, or they did not feel they had the time (they could not leave work or obtain child care) to attend the sessions. This group of non-participants may represent a small but meaningful subgroup of respondents. As well, although a previous MVC was a source of exclusion, a history of non-specific neck pain or neck injuries associated to other mechanisms (sports or falls) were not a source of exclusion and it is unknown if these would influence the outcomes. Unfortunately, adding these as exclusion criteria could severely limit the participant pool and would likely skew the average age to a much younger group thereby limiting the overall interpretation of the results. With limitations that are both unknown and unable to be controlled, the interpretation of the results will be limited as a cause and effect relationship cannot be stated and we are therefore restricted to recognizing correlational relationships.

Delimitations also needed to be managed in order to create the models based on these variables. One of the primary issues first encountered was how to categorize the participants into the two groups, recovered and persistent symptoms, this was an issue because it is not known at what percentage on the NDI someone is considered to have mild symptoms. In one study, recovered is said to be 10% and below; thus, mild is above 10% and below 30%.[77] In other studies the recovered group was below 10%, while mild included 10% to below 30%.[19] Evaluating which percentage cutoff to be considered mild or recovered may seem trivial, but in this study 12 participants had an NDI score of 10% at T4. Categorizing them as mild symptoms

resulted in 60 participants being assigned to the persistent symptoms group and 37 to the recovered group. If those same people had been considered recovered, there would have been 48 in the persistent symptoms group and 49 in the recovered group, and this may have altered the statistical analysis. However, since one of the goals of this study was to minimize false negatives, it was determined the best option was to set the mild symptoms at 10% for their final NDI score.

A statistical delimitation within this study is recognized by the fact that the sleep NDI score was used to predict future NDI scores, and this can lead to an inflation of the effect size of the variable due to the IV being an item response of the DV. However, two primary concepts kept this variable within the study. The first was related to a previous report that noted that a brief 5-item NDI was as statistically sound and responsive as the standard 10-item survey.[140] With the 5-item survey it was determined that the sleep response item was not necessary to determine a person's function. As a result, the sleep response score can be thought of as an independent measure not fully associated with the persons function as measured by the NDI.

Further justification for using the sleep NDI as an IV was strengthened when the correlation between the sleep NDI scores from T1 to T4 and the NDI T4 score were evaluated. The *r* values for the scores were T1 (0.28), T2 (0.19), T3 (0.50), and T4 (0.67). All *r* values except T2 had a significance level below 0.05. The correlation, although significant, ranged only from weak to modest.[136] With its limited strength as a correlation coefficient, as well as a previous report that noted that sleep NDI was not a necessary part of the functional assessment, it was considered appropriate to use the sleep response item as a variable of prediction.

Other delimitations include the use of ordinal data with sleep response scores and the NPRS. Although it was found to be appropriate to create a categorical variable for sleep response

scores and to use the NPRS as continuous data, these were still transformations of the variables that may have had an influence on the outcome.

Finally, in this study, we were not able to input data into the model from psychosocial stressors such as PTSD, depression, or anxiety, which have been found to have an influence on persistent symptoms.[82,103] Future studies should work to combine known psychosocial variables, such as PTSD, along with the variables reported in this study to determine the most robust model for predicting future outcomes.

5.7 Summary

WAD is not a new phenomenon, it has been a reported condition before automobiles were even manufactured due to injuries on the railway.[46] Nowadays, WAD following an MVC is a common condition seen in a physical therapy practice with an estimated 1 to 2.5 million cases occurring each year in the United States.[52,53] However, even with patient management and treatment strategies that are based on significant quantities of research and advanced diagnostics that attempt to understand WAD and its complexities, studies consistently demonstrate that approximately 50% of these patients will continue to have symptoms at least 6 to 12-months post-MVC.[77]

Some of the struggles with managing the patient with WAD are likely due to the intricate biopsychosocial interactions that affect each person and can result in a myriad of signs and symptoms including but not limited to neck pain, headaches, loss of range of motion, numbness, tingling, confusion, lethargy, insomnia, or difficulty concentrating.[2,3] The possibility that any or all of these may be present in any person, or be present in any number of combinations, makes it difficult for the practitioner to fully grasp the condition and be able to provide help to the patient.

However, before we can begin to treat these patients, we need to determine who will develop persistent symptoms. Currently, most of our clinical diagnostics have been focused on determining who will have moderate to severe symptoms 6 to 12-months post-MVC. In the derivation of a CPR, the variables of older age ($\geq$ 35 years old), elevated scores on the hyperarousal subscale of the post-traumatic stress distress scale ($\geq$ 6), or elevated NDI scores ($\geq$ 40) were the variables found to be the strongest predictors of the moderate to severe category.[77] In this CPR, they also determined who was likely to move to full recovery defined as an NDI of less than 10%. In this section of the study the authors reported that an NDI at or below 32 as well as age below 35 were predictors of full recovery. However, there were many pathways and combinations of the variables that did not lead to a recognition of who will be fully recovered or develop persistent symptoms.

Therefore, in this study, one of the primary aims was to determine who will develop persistent symptoms regardless if they were considered mild or moderate to severe. To achieve this goal, a two-group model of persistent symptoms and recovered was used; participants were categorized based on their NDI scores at 12 months with those below 10% labeled as recovered, while those with 10% or above were considered to have persistent symptoms.

When determining which variables to use as predictors in this study, the NDI was strongly considered based on its strength as a predictive variable for persistent symptoms. However, because the NDI is the tool used to determine group membership at 12 months, it was deemed inappropriate to use as an IV in the logistic regression model. Therefore, other variables that have been found to have at least some relationship with chronic WAD were evaluated in the model, and they were evaluated across multiple time points to determine if their strength of association changed. In this study the main effects for BMI, age, NPRS, sex, MFI, sleep NDI

were assessed. In addition to these main effects, interaction effects between variables that were different between the groups at baseline were also evaluated, and these included sex by MFI, sex by sleep NDI, and MFI by sleep NDI.

As this study was exploratory in nature, there was an emphasis on minimizing false negatives and ensuring all variables that were predictive would be included in the final model; thus, the parameters for inclusion were modified accordingly. Some of the adjustments to the statistical analysis included an elevated alpha level, reducing the level at which the score on the NDI was used to determine persistent symptoms (10% compared to 30%), and the use of the more lenient backwards conditional method for entering the variables in the logistic regression model. Therefore, if a variable was not found to be a significant predictor of group membership given these liberal standards, it was determined that they have little effect on group discrimination. From this model, the primary variables' main effects that were found to be predictors of persistent symptoms were MFI, sex, and the sleep NDI responses; no interactions were found to be significant at any of the time points.

MFI is a variable that has been found consistently in higher quantities in people who develop moderate to severe chronic WAD when compared to those who recover.[18,34] In previous literature, MFI has been found to be equal across groups at baseline and is reported to be in greater quantities in those who develop persistent moderate to severe symptoms in as little as 2 weeks and these differences maintain through at least 6 months.[19] Additionally, the reported effect size for MFI has been reported to range from medium to large, but this is often based on evaluating participants in the moderate to severe group when their NDI was only above 30%.[19,20]

In the results from this study, MFI was found to be different between the groups at baseline and this difference persisted across all time points, but it did not significantly across

time. Its effect size was also considered to be small, which is different from the medium and large effect sizes previously reported. Although these results differ from the previous literature, it may not be unanticipated based on the different categorization of the subjects. Regardless of the specific effect size, it appears that MFI is a consistent factor in the development of chronic WAD. In the future, MFI should be recognized as an important element in research that attempts to predict those who will develop persistent symptoms as well as in studies that evaluate best methods for treating people with the condition.

Sex is a variable that has not been found to be as consistent as MFI as a predictor of persistent symptoms,[17] but in this study, it was reported to have significant main effects across all four time points along with consistent effect sizes. This variable, although it is not amenable to changes during treatment, does provide some benefit to the treating clinician as they attempt to determine who will develop chronic WAD symptoms and may require treatment.

The sleep NDI response evaluates the amount of sleep disturbance a person is reporting. In the literature, there is some research and discussion about sleep disturbances as a predictor of persistent symptoms,[119,120] but its effect size has been reported as small and even these results are inconclusive.[17] However, in the past few years there has been an increase in research into the role of chronic sleep disturbances and its influence on chronic pain.[117] As our understanding of sleep disturbances and its bidirectional relationship with chronic pain is further developed, recognizing how it can be a part of the chronic WAD presentation will become clearer. With the results from this study, it appears that sleep disturbances function as an important predictor in developing chronic WAD, and whether it is a lot or just a little sleep loss, the odds of developing persistent symptoms increase. Additionally, the relationship between sleep loss and persistent symptoms appears to increase in strength as the time from the MVC increases for the person. Future studies

that evaluate those with WAD should consider sleep as a strong predictive variable that needs to be assessed, and it would be of interest to further delineate the subcategories of sleep disturbances (not just quantity of sleep loss) to understand how they may influence persistent symptoms.

These variables, although providing some information into the inner workings of WAD, are not definitive in their ability to predict group membership as the R^2 of the models listed ranged from 0.27 to 0.56. Thus, there appears to be a large amount of variance of the DV that is not explained by these variables. Additionally, the ability of these models to discriminate between the two groups was considered poor to acceptable, depending on the time point.

The results from this study demonstrate that some variables, sex and MFI, provide consistent effects across time on their influence on the development of chronic WAD, whereas sleep NDI appears to strengthen in its effect size as time progresses. However, the overall model prediction is still quite limited and needs further refinement to help discriminate between groups. Nonetheless, future studies that are attempting to predict who will develop persistent symptoms should include these variables in their assessment based on their consistency and effect sizes.

References

1. Gustafsson M, Stigson H, Krafft M, Kullgren A. Risk of permanent medical impairment (RPMI) in car crashes correlated to age and gender. *Traffic Inj Prev.* 2015;16(4):353-361.
2. Spitzer WO, Skovron ML, Salmi LR, et al. Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: redefining "whiplash" and its management. *Spine (Phila Pa 1976).* 1995;20(8 Suppl):1s-73s.
3. Smith AC, Parrish TB, Hoggarth MA, et al. Potential associations between chronic whiplash and incomplete spinal cord injury. *Spinal Cord Ser Cases.* 2015;1.
4. Ritchie C, Sterling M. Recovery Pathways and Prognosis After Whiplash Injury. *J Orthop Sports Phys Ther.* 2016;46(10):851-861.
5. Matsumoto M, Okada E, Ichihara D, et al. Prospective ten-year follow-up study comparing patients with whiplash-associated disorders and asymptomatic subjects using magnetic resonance imaging. *Spine (Phila Pa 1976).* 2010;35(18):1684-1690.
6. Lord SM, Barnsley L, Wallis BJ, Bogduk N. Chronic cervical zygapophysial joint pain after whiplash. A placebo-controlled prevalence study. *Spine (Phila Pa 1976).* 1996;21(15):1737-1744; discussion 1744-1735.
7. Elliott JM, Hancock MJ, Crawford RJ, Smith AC, Walton DM. Advancing imaging technologies for patients with spinal pain: with a focus on whiplash injury. *Spine J.* 2017.
8. Yoganandan N, Cusick JF, Pintar FA, Rao RD. Whiplash injury determination with conventional spine imaging and cryomicrotomy. *Spine (Phila Pa 1976).* 2001;26(22):2443-2448.
9. U.S Food and Drug Administration. What are the radiation risks from CT scan? 2009; https://www.fda.gov/Radiation-EmittingProducts/RadiationEmittingProductsandProcedures/MedicalImaging/MedicalX-Rays/ucm115329.htm. Accessed 9/27/2017.
10. Westbrook C, Roth CK. *MRI in Practice.* Chichester, West Sssex, UK: John Wiley & Sons; 2011.
11. Nakashima H, Yukawa Y, Suda K, Yamagata M, Ueta T, Kato F. Cervical Disc Protrusion Correlates With the Severity of Cervical Disc Degeneration: A Cross-Sectional Study of 1211 Relatively Healthy Volunteers. *Spine (Phila Pa 1976).* 2015;40(13):E774-779.
12. Elliott J, Sterling M, Noteboom JT, Darnell R, Galloway G, Jull G. Fatty infiltrate in the cervical extensor muscles is not a feature of chronic, insidious-onset neck pain. *Clin Radiol.* 2008;63(6):681-687.
13. Abbott R, Peolsson A, West J, et al. The Qualitative Grading of Muscle Fat Infiltration in Whiplash Using Fat/Water Magnetic Resonance Imaging. *Spine J.* 2017.
14. Krakenes J, Kaale BR. Magnetic resonance imaging assessment of craniovertebral ligaments and membranes after whiplash trauma. *Spine (Phila Pa 1976).* 2006;31(24):2820-2826.
15. Vangronsveld K, Peters M, Goossens M, Linton S, Vlaeyen J. Applying the fear-avoidance model to the chronic whiplash syndrome. *Pain.* 2007;131(3):258-261.
16. Ferrari R. A prospective study of perceived injustice in whiplash victims and its relationship to recovery. *Clin Rheumatol.* 2015;34(5):975-979.

17. Walton DM, Carroll LJ, Kasch H, et al. An Overview of Systematic Reviews on Prognostic Factors in Neck Pain: Results from the International Collaboration on Neck Pain (ICON) Project. *Open Orthop J.* 2013;7:494-505.
18. Elliott JM, O'Leary S, Sterling M, Hendrikz J, Pedler A, Jull G. Magnetic resonance imaging findings of fatty infiltrate in the cervical flexors in chronic whiplash. *Spine (Phila Pa 1976).* 2010;35(9):948-954.
19. Elliott J, Pedler A, Kenardy J, Galloway G, Jull G, Sterling M. The temporal development of fatty infiltrates in the neck muscles following whiplash injury: an association with pain and posttraumatic stress. *PloS One.* 2011;6(6):e21194.
20. Elliott JM, Courtney DM, Rademaker A, Pinto D, Sterling MM, Parrish TB. The Rapid and Progressive Degeneration of the Cervical Multifidus in Whiplash: An MRI Study of Fatty Infiltration. *Spine (Phila Pa 1976).* 2015;40(12):E694-700.
21. Davis CG. Mechanisms of chronic pain from whiplash injury. *J Forensic Leg Med.* 2013;20(2):74-85.
22. Bartsch AJ, Gilbertson LG, Prakash V, Morr DR, Wiechel JF. Minor Crashes and 'Whiplash' in the United States. *Ann Adv Automot Med.* 2008;52:117-130.
23. Ioppolo F, Rizzo R. Epidemiology of whiplash-associated disorders. *Whiplash Injuries*. Milano, Italy: Springer; 2014:13-16.
24. Anderson SE, Boesch C, Zimmermann H, et al. Are there cervical spine findings at MR imaging that are specific to acute symptomatic whiplash injury? A prospective controlled study with four experienced blinded readers. *Radiology.* 2012;262(2):567-575.
25. Ronnen HR, de Korte PJ, Brink PR, van der Bijl HJ, Tonino AJ, Franke CL. Acute whiplash injury: is there a role for MR imaging?--a prospective study of 100 patients. *Radiology.* 1996;201(1):93-96.
26. MacNab I. Acceleration injuries of the cervical spine. *J Bone Joint Surg Am.* 1964;46(8):1797-1799.
27. Barnsley L, Lord SM, Wallis BJ, Bogduk N. The prevalence of chronic cervical zygapophysial joint pain after whiplash. *Spine (Phila Pa 1976).* 1995;20(1):20-25; discussion 26.
28. Bogduk N. On cervical zygapophysial joint pain after whiplash. *Spine (Phila Pa 1976).* 2011;36(25 Suppl):S194-199.
29. Sterling M, McLean SA, Sullivan MJ, Elliott JM, Buitenhuis J, Kamper SJ. Potential processes involved in the initiation and maintenance of whiplash-associated disorders: discussion paper 3. *Spine (Phila Pa 1976).* 2011;36(25 Suppl):S322-329.
30. Walton DM, Elliott JM. An Integrated Model of Chronic Whiplash-Associated Disorder. *J Orthop Sports Phys Ther.* 2017;47(7):462-471.
31. Crawford RJ, Cornwall J, Abbott R, Elliott JM. Manually defining regions of interest when quantifying paravertebral muscles fatty infiltration from axial magnetic resonance imaging: a proposed method for the lumbar spine with anatomical cross-reference. *BMC Musculoskelet Disord.* 2017;18(1):25.
32. Addison O, Marcus RL, Lastayo PC, Ryan AS. Intermuscular fat: a review of the consequences and causes. *Int J Endocrinol.* 2014;2014:309570.
33. Sterling M, Jull G, Wright A. The effect of musculoskeletal pain on motor activity and control. *J Pain.* 2001;2(3):135-145.

34. Elliott J, Jull G, Noteboom JT, Darnell R, Galloway G, Gibbon WW. Fatty infiltration in the cervical extensor muscles in persistent whiplash-associated disorders: a magnetic resonance imaging analysis. *Spine (Phila Pa 1976).* 2006;31(22):E847-855.
35. Sterling M, Elliott JM, Cabot PJ. The course of serum inflammatory biomarkers following whiplash injury and their relationship to sensory and muscle measures: a longitudinal cohort study. *PloS One.* 2013;8(10):e77903.
36. Walton DM, Macdermid JC, Russell E, Koren G, Van Uum S. Hair-Normalized Cortisol Waking Response as a Novel Biomarker of Hypothalamic-Pituitary-Adrenal Axis Activity following Acute Trauma: A Proof-of-Concept Study with Pilot Results. *Pain Res Treat.* 2013;2013:876871.
37. Lund JP, Donga R, Widmer CG, Stohler CS. The pain-adaptation model: a discussion of the relationship between chronic musculoskeletal pain and motor activity. *Can J Physiol Pharmacol.* 1991;69(5):683-694.
38. Lethem J, Slade PD, Troup JD, Bentley G. Outline of a Fear-Avoidance Model of exaggerated pain perception--I. *Behav Res Ther.* 1983;21(4):401-408.
39. Zetterqvist V, Holmstrom L, Maathz P, Wicksell RK. Pain avoidance predicts disability and depressive symptoms three years later in individuals with whiplash complaints. *Acta Anaesthesiol Scand.* 2017;61(4):445-455.
40. McLean SA. The potential contribution of stress systems to the transition to chronic whiplash-associated disorders. *Spine (Phila Pa 1976).* 2011;36(25 Suppl):S226-232.
41. Passatore M, Roatta S. Influence of sympathetic nervous system on sensorimotor function: whiplash associated disorders (WAD) as a model. *Eur J Appl Physiol.* 2006;98(5):423-449.
42. Todman D. Whiplash injuries: A historical review. *Internet J Neurol.* 2007;8(2).
43. Cote P, Wong JJ, Sutton D, et al. Management of neck pain and associated disorders: A clinical practice guideline from the Ontario Protocol for Traffic Injury Management (OPTIMa) Collaboration. *Eur Spine J.* 2016;25(7):2000-2022.
44. Jull GA. Considerations in the physical rehabilitation of patients with whiplash-associated disorders. *Spine (Phila Pa 1976).* 2011;36(25 Suppl):S286-291.
45. Norris SH, Watt I. The prognosis of neck injuries resulting from rear-end vehicle collisions. *J Bone Joint Surg Br.* 1983;65(5):608-611.
46. Harrington R. On the tracks of trauma: railway spine reconsidered. *Soc Hist Med.* 2003;16(2):209-223.
47. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5®).* 5th ed. Arlington, VA: American Psychiatric Pub; 2013.
48. MacDermid JC, Walton DM, Avery S, et al. Measurement properties of the neck disability index: a systematic review. *J Orthop Sports Phys Ther.* 2009;39(5):400-417.
49. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess.* 1995;7(4):524.
50. Foa EB, Cashman L, Jaycox L, Perry K. The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. *Psychol Assess.* 1997;9(4):445.
51. Peever J, Fuller PM. The Biology of REM Sleep. *Curr Biol.* 2017;27(22):R1237-r1248.
52. Haneline MT. Epidemiology of Whiplash Injuries. *Whiplash: A Patient Centered Approach To Management*. St. Louis, MO: Elsevier; 2012:154-168.

53. Centers for Disease Control and Prevention. Motor Vehicle Crash Injuries. 2014; https://www.cdc.gov/vitalsigns/crash-injuries/index.html. Accessed 6/13/2018.
54. National Safety Council. Injury Facts Motor Vehicle Overview. 2020; https://injuryfacts.nsc.org/motor-vehicle/overview/introduction/. Accessed 2/4/2020.
55. Freeman MD, Croft AC, Rossignol AM, Weaver DS, Reiser M. A review and methodologic critique of the literature refuting whiplash syndrome. *Spine (Phila Pa 1976).* 1999;24(1):86-96.
56. Pink J, Petrou S, Williamson E, Williams M, Lamb SE. Economic and Health-Related Quality of Life Outcomes of Whiplash Associated Disorders. *Spine (Phila Pa 1976).* 2016;41(17):1378-1386.
57. Siegmund G. The biomechanics of whiplash injury. *B C Med J.* 2002;44(5):243-247.
58. Siegmund GP. What occupant kinematics and neuromuscular responses tell us about whiplash injury. *Spine (Phila Pa 1976).* 2011;36(25 Suppl):S175-179.
59. Stemper BD, Corner BD. Whiplash-Associated Disorders: Occupant Kinematics and Neck Morphology. *J Orthop Sports Phys Ther.* 2016;46(10):834-844.
60. Ivancic PC, Ito S, Tominaga Y, et al. Whiplash causes increased laxity of cervical capsular ligament. *Clin Biomech (Bristol, Avon).* 2008;23(2):159-165.
61. Hendey GW, Votey SR. Injuries in restrained motor vehicle accident victims. *Ann Emerg Med.* 1994;24(1):77-84.
62. Uhrenholt L, Grunnet-Nilsson N, Hartvigsen J. Cervical spine lesions after road traffic accidents: a systematic review. *Spine (Phila Pa 1976).* 2002;27(17):1934-1941; discussion 1940.
63. Malleson A. Chronic whiplash syndrome. Psychosocial epidemic. *Can Fam Physician.* 1994;40:1906-1909.
64. Kasch H, Turk DC, Jensen TS. *Whiplash Injury: Perspectives on the Development of Chronic Pain.* Washington, D.C: Wolters Kluwer Health; 2016.
65. Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. *Pain.* 2011;152(3 Suppl):S2-15.
66. Ferrari R. *The whiplash encyclopedia: the facts and myths of whiplash.* Burlington, MA: Jones & Bartlett Learning; 1999.
67. Sterling M, Carroll LJ, Kasch H, Kamper SJ, Stemper B. Prognosis after whiplash injury: where to from here? Discussion paper 4. *Spine (Phila Pa 1976).* 2011;36(25 Suppl):S330-334.
68. Sterling M, Jull G, Kenardy J. Physical and psychological factors maintain long-term predictive capacity post-whiplash injury. *Pain.* 2006;122(1-2):102-108.
69. Casey PP, Feyer AM, Cameron ID. Associations with duration of compensation following whiplash sustained in a motor vehicle crash. *Injury.* 2015;46(9):1848-1855.
70. Gentle CR, Golinski WZ, Heitplatz F. Computational studies of 'whiplash' injuries. *Proc Inst Mech Eng H.* 2001;215(2):181-189.
71. Bogduk N, Yoganandan N. Biomechanics of the cervical spine Part 3: minor injuries. *Clin Biomech (Bristol, Avon).* 2001;16(4):267-275.
72. Bogduk N. *Chapter 53 Whiplash injury.* Vol 81. 2008/09/24 ed. St. Louis, MO: Elsevier; 2006.
73. Bogduk N, Mercer S. Biomechanics of the cervical spine. I: Normal kinematics. *Clin Biomech (Bristol, Avon).* 2000;15(9):633-648.

74. Knight K, Knight KL, Draper D. *Therapeutic modalities: the art and science.* 2nd ed. Baltimore, MD: Lippincott Williams & Wilkins; 2012.
75. Hartling L, Brison RJ, Ardern C, Pickett W. Prognostic value of the Quebec Classification of Whiplash-Associated Disorders. *Spine (Phila Pa 1976).* 2001;26(1):36-41.
76. Kivioja J, Jensen I, Lindgren U. Neither the WAD-classification nor the Quebec Task Force follow-up regimen seems to be important for the outcome after a whiplash injury. A prospective study on 186 consecutive patients. *Eur Spine J.* 2008;17(7):930-935.
77. Ritchie C, Hendrikz J, Kenardy J, Sterling M. Derivation of a clinical prediction rule to identify both chronic moderate/severe disability and full recovery following whiplash injury. *Pain.* 2013;154(10):2198-2206.
78. Ritchie C, Hendrikz J, Jull G, Elliott J, Sterling M. External validation of a clinical prediction rule to predict full recovery and ongoing moderate/severe disability following acute whiplash injury. *J Orthop Sports Phys Ther.* 2015;45(4):242-250.
79. McLean SA, Clauw DJ, Abelson JL, Liberzon I. The development of persistent pain and psychological morbidity after motor vehicle collision: integrating the potential role of stress response systems into a biopsychosocial model. *Psychosom Med.* 2005;67(5):783-790.
80. Curatolo M, Bogduk N, Ivancic PC, McLean SA, Siegmund GP, Winkelstein BA. The role of tissue damage in whiplash-associated disorders: discussion paper 1. *Spine (Phila Pa 1976).* 2011;36(25 Suppl):S309-315.
81. Coughlin TR, Kennedy OD. The role of subchondral bone damage in post-traumatic osteoarthritis. *Ann N Y Acad Sci.* 2016;1383(1):58-66.
82. Walton DM, Macdermid JC, Giorgianni AA, Mascarenhas JC, West SC, Zammit CA. Risk factors for persistent problems following acute whiplash injury: update of a systematic review and meta-analysis. *J Orthop Sports Phys Ther.* 2013;43(2):31-43.
83. Linnstaedt SD, Riker KD, Walker MG, et al. MicroRNA 320a Predicts Chronic Axial and Widespread Pain Development Following Motor Vehicle Collision in a Stress-Dependent Manner. *J Orthop Sports Phys Ther.* 2016;46(10):911-919.
84. Melzack R, Katz J. Pain. *Wiley Interdiscip Rev Cogn Sci.* 2013;4(1):1-15.
85. Tuttle LJ, Sinacore DR, Mueller MJ. Intermuscular adipose tissue is muscle specific and associated with poor functional performance. *J Aging Res.* 2012;2012:172957.
86. Delmonico MJ, Harris TB, Visser M, et al. Longitudinal study of muscle strength, quality, and adipose tissue infiltration. *Am J Clin Nutr.* 2009;90(6):1579-1585.
87. Goodpaster BH, Chomentowski P, Ward BK, et al. Effects of physical activity on strength and skeletal muscle fat infiltration in older adults: a randomized controlled trial. *J Appl Physiol (1985).* 2008;105(5):1498-1503.
88. Steilen D, Hauser R, Woldin B, Sawyer S. Chronic neck pain: making the connection between capsular ligament laxity and cervical instability. *Open Orthop J.* 2014;8:326-345.
89. Marcell TJ. Sarcopenia: causes, consequences, and preventions. *J Gerontol A Biol Sci Med Sci.* 2003;58(10):M911-916.
90. Gardner MP, Lightman S, Sayer AA, et al. Dysregulation of the hypothalamic pituitary adrenal (HPA) axis and physical performance at older ages: an individual participant meta-analysis. *Psychoneuroendocrinology.* 2013;38(1):40-49.

91. Scicchitano BM, Sica G, Musaro A. Stem Cells and Tissue Niche: Two Faces of the Same Coin of Muscle Regeneration. *Eur J Transl Myol.* 2016;26(4):6125.
92. Hamrick MW, McGee-Lawrence ME, Frechette DM. Fatty Infiltration of Skeletal Muscle: Mechanisms and Comparisons with Bone Marrow Adiposity. *Front Endocrinol (Lausanne).* 2016;7:69.
93. Sciorati C, Clementi E, Manfredi AA, Rovere-Querini P. Fat deposition and accumulation in the damaged and inflamed skeletal muscle: cellular and molecular players. *Cell Mol Life Sci.* 2015;72(11):2135-2156.
94. Abdelhamid RE, Sluka KA. ASICs Mediate Pain and Inflammation in Musculoskeletal Diseases. *Physiology (Bethesda).* 2015;30(6):449-459.
95. Elliott J, Sterling M, Noteboom JT, Treleaven J, Galloway G, Jull G. The clinical presentation of chronic whiplash and the relationship to findings of MRI fatty infiltrates in the cervical extensor musculature: a preliminary investigation. *Eur Spine J.* 2009;18(9):1371-1378.
96. O'Leary S, Jull G, Van Wyk L, Pedler A, Elliott J. Morphological changes in the cervical muscles of women with chronic whiplash can be modified with exercise-A pilot study. *Muscle Nerve.* 2015;52(5):772-779.
97. Brinjikji W, Luetmer PH, Comstock B, et al. Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. *AJNR Am J Neuroradiol.* 2015;36(4):811-816.
98. Volpi E, Nazemi R, Fujita S. Muscle tissue changes with aging. *Curr Opin Clin Nutr Metab Care.* 2004;7(4):405-410.
99. Tsang A, Von Korff M, Lee S, et al. Common chronic pain conditions in developed and developing countries: gender and age differences and comorbidity with depression-anxiety disorders. *The journal of pain.* 2008;9(10):883-891.
100. Griggs RK, Cook J, Gargan M, Bannister G, Amirfeyz R. Mid-term follow-up of whiplash with Bournemouth Questionnaire: the significance of the initial depression to pain ratio. *J Back Musculoskelet Rehabil.* 2015;28(3):443-446.
101. Gargan M, Bannister G, Main C, Hollis S. The behavioural response to whiplash injury. *J Bone Joint Surg Br.* 1997;79(4):523-526.
102. Lankester BJ, Garneti N, Gargan MF, Bannister GC. Factors predicting outcome after whiplash injury in subjects pursuing litigation. *Eur Spine J.* 2006;15(6):902-907.
103. Sterling M, Hendrikz J, Kenardy J. Similar factors predict disability and posttraumatic stress disorder trajectories after whiplash injury. *Pain.* 2011;152(6):1272-1278.
104. Kitamura T, Nakagawa Y, Machizawa S. Grading depression severity by symptom scores: is it a valid method for subclassifying depressive disorders? *Compr Psychiatry.* 1993;34(4):280-283.
105. Vecchio PC. Is whiplash real? *The Medical Journal of Australia.* 2003;180(2).
106. Hadler NM. If you have to prove you are ill, you can't get well. The object lesson of fibromyalgia. *Spine (Phila Pa 1976).* 1996;21(20):2397-2400.
107. Haneline MT. The notion of a “whiplash culture”: a review of the evidence. *J Chiropr Med.* 2009;8(3):119-124.
108. Carroll LJ, Holm LW, Hogg-Johnson S, et al. Course and prognostic factors for neck pain in whiplash-associated disorders (WAD): results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *J Manipulative Physiol Ther.* 2009;32(2 Suppl):S97-s107.

109. Obelieniene D, Schrader H, Bovim G, Misevičiene I, Sand T. Pain after whiplash: a prospective controlled inception cohort study. *Journal of Neurology, Neurosurgery & Psychiatry.* 1999;66(3):279-283.
110. Schrader H, Obelieniene D, Bovim G, et al. Natural evolution of late whiplash syndrome outside the medicolegal context. *Lancet.* 1996;347(9010):1207-1211.
111. Cohen SP. Epidemiology, diagnosis, and treatment of neck pain. *Mayo Clin Proc.* 2015;90(2):284-299.
112. Viano DC. Seat influences on female neck responses in rear crashes: a reason why women have higher whiplash rates. *Traffic Inj Prev.* 2003;4(3):228-239.
113. Beckles GL, Truman BI, Control CfD, Prevention. Education and income—United States, 2005 and 2009. *MMWR Surveill Summ.* 2011;60(Suppl):13-17.
114. Sejnowski TJ, Destexhe A. Why do we sleep? *Brain Res.* 2000;886(1-2):208-223.
115. Finan PH, Goodin BR, Smith MT. The association of sleep and pain: an update and a path forward. *J Pain.* 2013;14(12):1539-1552.
116. Mullington JM, Simpson NS, Meier-Ewert HK, Haack M. Sleep loss and inflammation. *Best Pract Res Clin Endocrinol Metab.* 2010;24(5):775-784.
117. Nijs J, Mairesse O, Neu D, et al. Sleep Disturbances in Chronic Pain: Neurobiology, Assessment, and Treatment in Physical Therapist Practice. *Phys Ther.* 2018;98(5):325-335.
118. Curatolo M, Möller M, Ashraf A, et al. Pain hypersensitivity and spinal nociceptive hypersensitivity in chronic pain: prevalence and associated factors. *Pain (03043959).* 2015;156(11):2373-2382.
119. Hendriks EJ, Scholten-Peeters GG, van der Windt DA, Neeleman-van der Steen CW, Oostendorp RA, Verhagen AP. Prognostic factors for poor recovery in acute whiplash patients. *Pain.* 2005;114(3):408-416.
120. Valenza MC, Valenza G, Gonzalez-Jimenez E, De-la-Llave-Rincon AI, Arroyo-Morales M, Fernandez-de-Las-Penas C. Alteration in sleep quality in patients with mechanical insidious neck pain and whiplash-associated neck pain. *Am J Phys Med Rehabil.* 2012;91(7):584-591.
121. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine (Phila Pa 1976).* 2001;26(17):1873-1878.
122. Hancock MJ, Maher CM, Petocz P, et al. Risk factors for a recurrence of low back pain. *Spine J.* 2015;15(11):2360-2368.
123. Hill DL, Batchelor PG, Holden M, Hawkes DJ. Medical image registration. *Phys Med Biol.* 2001;46(3):R1-45.
124. Sterling M, Hendrikz J, Kenardy J. Compensation claim lodgement and health outcome developmental trajectories following whiplash injury: A prospective study. *Pain.* 2010;150(1):22-28.
125. Walton DM, Elliott JM. A new clinical model for facilitating the development of pattern recognition skills in clinical pain assessment. *Musculoskelet Sci Pract.* 2018;36:17-24.
126. Meeus M, Nijs J, Hamers V, Ickmans K, Oosterwijck JV. The efficacy of patient education in whiplash associated disorders: a systematic review. *Pain Physician.* 2012;15(5):351-361.

127. Elliott JM, Walton DM, Rademaker A, Parrish TB. Quantification of cervical spine muscle fat: a comparison between T1-weighted and multi-echo gradient echo imaging using a variable projection algorithm (VARPRO). *BMC Med Imaging.* 2013;13:30.
128. Ogawa M, Lester R, Akima H, Gorgey AS. Quantification of intermuscular and intramuscular adipose tissue using magnetic resonance imaging after neurodegenerative disorders. *Neural Regen Res.* 2017;12(12):2100-2105.
129. Casey PP, Feyer AM, Cameron ID. Course of recovery for whiplash associated disorders in a compensation setting. *Injury.* 2015;46(11):2118-2129.
130. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther.* 1991;14(7):409-415.
131. Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy. *Spine (Phila Pa 1976).* 2006;31(5):598-602.
132. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. *J Manipulative Physiol Ther.* 2008;31(7):491-502.
133. Nieto R, Miro J, Huguet A. Disability in subacute whiplash patients: usefulness of the neck disability index. *Spine (Phila Pa 1976).* 2008;33(18):E630-635.
134. Hoving JL, O'Leary EF, Niere KR, Green S, Buchbinder R. Validity of the neck disability index, Northwick Park neck pain questionnaire, and problem elicitation technique for measuring disability associated with whiplash-associated disorders. *Pain.* 2003;102(3):273-281.
135. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53.
136. Portney L, Watkins M. *Foundations of Clinical Research Applications to Practice.* 3rd ed. Upper saddle River, NJ: Pearson Education; 2009.
137. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil.* 2008;89(1):69-74.
138. Vos CJ, Verhagen AP, Koes BW. Reliability and responsiveness of the Dutch version of the Neck Disability Index in patients with acute neck pain in general practice. *Eur Spine J.* 2006;15(11):1729-1736.
139. Pool JJ, Ostelo RW, Hoving JL, Bouter LM, de Vet HC. Minimal clinically important change of the Neck Disability Index and the Numerical Rating Scale for patients with neck pain. *Spine (Phila Pa 1976).* 2007;32(26):3047-3051.
140. Walton DM, MacDermid JC. A brief 5-item version of the Neck Disability Index shows good psychometric properties. *Health Qual Life Outcomes.* 2013;11:108.
141. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs.* 2005;14(7):798-804.
142. Bijur PE, Latimer CT, Gallagher EJ. Validation of a verbally administered numerical rating scale of acute pain for use in the emergency department. *Acad Emerg Med.* 2003;10(4):390-392.
143. Meyers L, Gamst G, Guarino A. *Applied Multivariate Research Design and Interpretation.* 2nd ed. Los Angeles, CA: SAGE; 2013.
144. Hosmer Jr DW, Lemeshow S, Sturdivant RX. *Applied logistic regression.* Vol 398: John Wiley & Sons; 2013.
145. Razali NM, Wah YB. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *J Stat Model Anal.* 2011;2(1):21-33.

146. Fisher AD, Ristori J, Morelli G, Maggi M. The molecular mechanisms of sexual orientation and gender identity. *Mol Cell Endocrinol.* 2018;467:3-13.
147. Ivancic PC. Mechanisms and Mitigation of Head and Spinal Injuries Due to Motor Vehicle Crashes. *J Orthop Sports Phys Ther.* 2016;46(10):826-833.
148. Kim H-Y. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restor Dent Endod.* 2013;38(1):52-54.
149. Kim TK. Practical statistics in pain research. *Korean J Pain.* 2017;30(4):243-249.
150. Tukey JW. The future of data analysis. *Ann Math Stat.* 1962;33(1):1-67.
151. Walton DM, Pretty J, MacDermid JC, Teasell RW. Risk factors for persistent problems following whiplash injury: results of a systematic review and meta-analysis. *J Orthop Sports Phys Ther.* 2009;39(5):334-350.
152. Kamper SJ, Rebbeck TJ, Maher CG, McAuley JH, Sterling M. Course and prognostic factors of whiplash: a systematic review and meta-analysis. *Pain.* 2008;138(3):617-629.
153. Williams M, Williamson E, Gates S, Lamb S, Cooke M. A systematic literature review of physical prognostic factors for the development of Late Whiplash Syndrome. *Spine (Phila Pa 1976).* 2007;32(25):E764-780.
154. Jeanmonod R, Varacallo M. Geriatric Cervical Spine Injury. *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2019.
155. Okada E, Matsumoto M, Ichihara D, et al. Aging of the cervical spine in healthy volunteers: a 10-year longitudinal magnetic resonance imaging study. *Spine (Phila Pa 1976).* 2009;34(7):706-712.
156. Hogg-Johnson S, van der Velde G, Carroll LJ, et al. The burden and determinants of neck pain in the general population: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Spine (Phila Pa 1976).* 2008;33(4 Suppl):S39-51.
157. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological reviews of the Cambridge Philosophical Society.* 2007;82(4):591-605.
158. Merskey H BN. *Classification of chronic pain: desription of chronic pain syndromes and definitions of pain terms.* Seattle, WA: ISAP Press; 1994.
159. U.S. Department of Justice. A Guide to Disability Rights Laws. In: Civil Rights Division, ed. Washington, D.C.2009.
160. Walton D. A review of the definitions of 'recovery' used in prognostic studies on whiplash using an ICF framework. *Disabil Rehabil.* 2009;31(12):943-957.
161. Norvell DC, Dettori JR, Chapman JR. Enhancing Clinical Study Retention Rates to Avoid Follow-up Bias: How Do We Keep Our Study Participants from "The Land of the Lost"? *Global spine journal.* 2016;6(5):519-521.
162. Elliott JM, Pedler AR, Jull GA, Van Wyk L, Galloway GG, O'Leary SP. Differential changes in muscle composition exist in traumatic and nontraumatic neck pain. *Spine (Phila Pa 1976).* 2014;39(1):39-47.
163. Sterling M, Jull G, Vicenzino B, Kenardy J, Darnell R. Development of motor system dysfunction following whiplash injury. *Pain.* 2003;103(1-2):65-73.
164. Falla D, Bilenkij G, Jull G. Patients with chronic neck pain demonstrate altered patterns of muscle activation during performance of a functional upper limb task. *Spine (Phila Pa 1976).* 2004;29(13):1436-1440.

165. Stenneberg MS, Rood M, de Bie R, Schmitt MA, Cattrysse E, Scholten-Peeters GG. To What Degree Does Active Cervical Range of Motion Differ Between Patients With Neck Pain, Patients With Whiplash, and Those Without Neck Pain? A Systematic Review and Meta-Analysis. *Arch Phys Med Rehabil.* 2017;98(7):1407-1434.
166. Woodhouse A, Vasseljen O. Altered motor control patterns in whiplash and chronic neck pain. *BMC Musculoskelet Disord.* 2008;9:90.
167. Anderson JS, Hsu AW, Vasavada AN. Morphology, architecture, and biomechanics of human cervical multifidus. *Spine (Phila Pa 1976).* 2005;30(4):E86-91.
168. Ackland DC, Merritt JS, Pandy MG. Moment arms of the human neck muscles in flexion, bending and rotation. *Journal of biomechanics.* 2011;44(3):475-486.
169. Vasavada AN, Li S, Delp SL. Influence of muscle morphometry and moment arms on the moment-generating capacity of human neck muscles. *Spine (Phila Pa 1976).* 1998;23(4):412-422.
170. Green CA, Pope CR. Gender, psychosocial factors and the use of medical services: a longitudinal analysis. *Social science & medicine (1982).* 1999;48(10):1363-1372.
171. Tudiver F, Talbot Y. Why don't men seek help? Family physicians' perspectives on help-seeking behavior in men. *The Journal of family practice.* 1999;48(1):47-52.
172. Chapline JF, Ferguson SA, Lillis RP, Lund AK, Williams AF. Neck pain and head restraint position relative to the driver's head in rear-end collisions. *Accid Anal Prev.* 2000;32(2):287-297.
173. Young AL, Ragel BT, Su E, Mann CN, Frank EH. Assessing automobile head restraint positioning in Portland, Oregon. *Inj Prev.* 2005;11(2):97-101.
174. Chouchou F, Khoury S, Chauny JM, Denis R, Lavigne GJ. Postoperative sleep disruptions: a potential catalyst of acute pain? *Sleep medicine reviews.* 2014;18(3):273-282.
175. Lautenbacher S, Kundermann B, Krieg JC. Sleep deprivation and pain perception. *Sleep medicine reviews.* 2006;10(5):357-369.
176. Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research.* 1989;28(2):193-213.
177. Namas R, Ghuma A, Hermus L, et al. The acute inflammatory response in trauma / hemorrhage and traumatic brain injury: current state and emerging prospects. *The Libyan journal of medicine.* 2009;4(3):97-103.
178. Chen L, Deng H, Cui H, et al. Inflammatory responses and inflammation-associated diseases in organs. *Oncotarget.* 2018;9(6):7204-7218.
179. Ozcan-Eksi EE, Eksi MS, Akcal MA. Severe Lumbar Intervertebral Disc Degeneration Is Associated with Modic Changes and Fatty Infiltration in the Paraspinal Muscles at all Lumbar Levels, Except for L1-L2: A Cross-Sectional Analysis of 50 Symptomatic Women and 50 Age-Matched Symptomatic Men. *World Neurosurg.* 2019;122:e1069-e1077.
180. Sasaki T, Yoshimura N, Hashizume H, et al. MRI-defined paraspinal muscle morphology in Japanese population: The Wakayama Spine Study. *PloS One.* 2017;12(11):e0187765.
181. de Boer M, Nijdam MJ, Jongedijk RA, et al. The Spectral Fingerprint of Sleep Problems in Post-Traumatic Stress Disorder. *Sleep.* 2019.

182. Weber KA, Smith AC, Wasielewski M, et al. Deep Learning Convolutional Neural Networks for the Automatic Quantification of Muscle Fat Infiltration Following Whiplash Injury. *Sci Rep.* 2019;9(1):7973.
183. Khoury V, Cardinal E, Brassard P. Atrophy and fatty infiltration of the supraspinatus muscle: sonography versus MRI. *AJR American journal of roentgenology.* 2008;190(4):1105-1111.

Appendices

Appendix A: Institutional Review Board

NOVA SOUTHEASTERN UNIVERSITY
Institutional Review Board

MEMORANDUM

To: **Brad Callan**

From: **Wendy Stav, PhD,**
Center Representative, Institutional Review Board

Date: **July 16, 2019**

Re: **IRB #: 2019-378; Title, "Neck Muscle Composition in Persistent Whiplash Associated Disorder: Relationship with Neck-Related Disability"**

I have reviewed the above-referenced research protocol at the center level. Based on the information provided, I have determined that this study is exempt from further IRB review under **45 CFR 46.101(b) (Exempt 4: Use of previously-collected records, data, specimens, tissues, etc.)**. You may proceed with your study as described to the IRB. As principal investigator, you must adhere to the following requirements:

1) CONSENT: If recruitment procedures include consent forms, they must be obtained in such a manner that they are clearly understood by the subjects and the process affords subjects the opportunity to ask questions, obtain detailed answers from those directly involved in the research, and have sufficient time to consider their participation after they have been provided this information. The subjects must be given a copy of the signed consent document, and a copy must be placed in a secure file separate from de-identified participant information. Record of informed consent must be retained for a minimum of three years from the conclusion of the study.

2) ADVERSE EVENTS/UNANTICIPATED PROBLEMS: The principal investigator is required to notify the IRB chair and me (954-262-5369 and Wendy Stav, PhD, respectively) of any adverse reactions or unanticipated events that may develop as a result of this study. Reactions or events may include, but are not limited to, injury, depression as a result of participation in the study, life-threatening situation, death, or loss of confidentiality/anonymity of subject. Approval may be withdrawn if the problem is serious.

3) AMENDMENTS: Any changes in the study (e.g., procedures, number or types of subjects, consent forms, investigators, etc.) must be approved by the IRB prior to implementation. Please be advised that changes in a study may require further review depending on the nature of the change. Please contact me with any questions regarding amendments or changes to your study.

The NSU IRB is in compliance with the requirements for the protection of human subjects prescribed in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) revised June 18, 1991.

Cc: Joshua Cleland, PT, PhD
Rose M Colon, PhD

Appendix B: Data Sharing Agreement

DATA USE AGREEMENT

This Data Use Agreement (**"Agreement"**) is effective as of May 28, 2019 (**"Effective Date"**) by and between Northwestern University (**"Provider"**), with principal offices located at 633 Clark Street, Evanston, IL 60208 and Nova Southeastern University (**"Recipient"**), a non-profit educational and research institution with principal offices located at 3301 College Ave., Fort Lauderdale, FL 33314 (collectively, the **"Parties"**).

WHEREAS, Provider maintains or shall obtain certain de-identified clinical data and information including Excel spreadsheet with the following: • A deidentified participant number. • Demographic data: age, gender, height, weight. • Quantities of muscle fat within the cervical musculature as measured by MRI at each of 4 time points. • Outcome scores at each of the 4 time points for the neck disability index, numeric pain rating scale, and sleep scores on the post-traumatic disability scale (**"Data"**) that Recipient wishes to use and/or disclose for research, public health, or health care operations purposes permitted under 45 C.F.R. §164.514(e) of regulations promulgated under HIPAA;

WHEREAS, with prior approval from the Provider, Recipient will receive the Data which from the Provider:

WHEREAS, Recipient desires to use Data for the research project entitled "Neck Muscle Composition in Persistent Whiplash Associated Disorder: Relationship with Neck Related Disability" (**"Study"**) under the supervision of Recipient investigator, Brad Callan, (**"Investigator"**).

The Parties agree to the provisions of this Agreement in order to address the applicable requirements of HIPAA and to protect the interest of both Parties.

1. RESTRICTIONS ON USE.

a. Data will be used solely for research purposes in connection with the Study by the Investigator and those employees of Recipient under his/her direct supervision ("Authorized Personnel"). Recipient agrees that it will not use or further disclose the Data other than as permitted by this Agreement, or as otherwise required by law or regulation. Recipient shall ensure that its Investigator and employees comply with the terms and conditions of this Agreement. Except as authorized under this Agreement or otherwise required by law, Recipient agrees to retain control over the Data and shall not disclose, release, sell, rent, lease, loan, or otherwise grant access to the Data to any third party, except Authorized Persons and then to the extent only for the purpose of this Agreement, without the prior written consent of Provider.

b. Recipient agrees to comply with all applicable federal, state and local laws and regulations applicable to the Study and relating to the maintenance of the Data and the use and disclosure of the Data. Additionally, Recipient agrees to comply with the institutional policies of Recipient as well as the policies set forth by its Institutional Review Board (IRB) in regards to Investigator's receipt and use of the Data. By signing this Agreement, Recipient provides assurance that relevant institutional policies and applicable federal, state, or local laws and regulations (if any) have been followed, including the completion of any IRB or ethics review or approval that may be required.

c. Recipient shall not use the Data, either alone or in concert with any other information, to make any effort to identify or contact individuals who are or may be sources of the Data without specific written

approval from Provider and appropriate Institutional Review Board (IRB) approval, if required pursuant to 45 CFR 46. Should Recipient inadvertently receive identifiable information or otherwise identify a subject, Recipient shall promptly notify Provider and follow Provider's reasonable written instructions, which may include return or destruction of the identifiable information

d. Recipient shall use appropriate safeguards to protect the Data from misuse or inappropriate disclosure and to prevent any use or disclosure of the Data other than as provided in this Agreement or as otherwise required by law or regulation. If Recipient becomes aware of a use or disclosure of Data by Recipient in violation of the requirements of this Agreement, it shall promptly notify Provider and take appropriate action to mitigate the harmful effects of such violation.

e. This Data represents a significant investment on the part of Provider and Provider shall retain ownership of the Data. The Provider represents that the Data are provided to Recipient in accordance with all applicable laws and regulations and in accordance with the terms of any consent document(s) under which they were obtained.

2. AUDITS

As reasonably necessary, Recipient agrees to make its internal practices and records, including policies and procedures, relating to the use and disclosure of Data available to Provider at a mutually agreed upon time during normal business hours upon prior written notice, provided that Provider shall be subject to reasonable safeguards to protect the security and integrity of Recipient's research enterprise and records systems.

3. PUBLICATIONS

Provider hereby grants to Recipient the right to use the Data for the limited purpose of carrying-out the Study, including publication resulting from the Study, provided that, in any publication, Recipient shall only use aggregated data derived from the Data and any disclosures of data and results derived from the Data must not contain the name of an individual or other information by which an individual can be directly or indirectly identified. Provider and Recipient agree to collaborate to make the results of the analysis of the Data publicly available by way of a joint publication or other joint disclosure. In any event, before Recipient or Investigator submits a paper or abstract for publication or otherwise publicly discloses information related to the Data, Recipient and Investigator shall ensure that Provider has at least thirty (30) days to review the proposed publication or disclosure. In all oral presentations or written publications concerning the Data, Recipient will acknowledge Provider's contribution of the Data, including naming the Provider's scientist as co-author, as scientifically appropriate, unless requested otherwise.

4. INTELLECTUAL PROPERTY.

Recipient agrees that if the Study results in an invention, a new use, or a product (collectively referred to as an "**Invention**") based on or relating to the Data, Recipient will promptly disclose the Invention to Provider on a confidential basis. Inventorship will be determined in accordance with U.S. patent law (if patentable) or by mutual agreement between the Parties (if not patentable), taking into account the role and contributions of individuals involved in the development of the Invention.

5. WAIVER OF WARRANTY; INDEMNIFICATION; INJUNCTIONS.

PROVIDER MAKES NO WARRANTIES, EXPRESS OR IMPLIED, AS TO ANY MATTER WHATSOEVER, INCLUDING, WITHOUT LIMITATION, THE DATA OR THE OWNERSHIP, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE OF SUCH DATA, OR ANY PRODUCT OR PROCESS BASED THEREON.

Except to the extent prohibited by law, the Recipient assumes all liability for damages which may arise from its use, storage or disposal of the Data. Recipient further agrees to indemnify, defend and hold harmless Provider, its trustees, directors, medical affiliates, employees, agents and subcontractors from and against any claims, damages, liabilities, or causes of actions (collectively **"Claims"**) arising out of or due to Recipient's negligence or willful misconduct.

Recipient agrees that any violation by Recipient of any of the provisions of this Agreement may cause irreparable harm to Provider. Accordingly, in addition to any other remedies available to Provider at law, in equity, or under this Agreement, Provider shall be entitled to seek an injunction or other decree of specific performance with respect to any violation by Recipient of any of the provisions of this Agreement, or any explicit threat thereof.

6. NOTICES.

Any notice required by this Agreement shall be made in writing and sent by U.S. mail, postage pre-paid, or nationally recognized overnight courier, to the addresses specified below:

If to Recipient:
Brad Callan
Nova Southeastern University
301 E. Mountainview Ave
Ellensburg, WA 98926

If to Provider:
James M. Elliott, PT, PhD
Northwestern University
Feinberg School of Medicine
Chicago IL 60611

For administrative matters:
Gary Margules, Sc.D.
Vice President for Office of Research,
and Technology Transfer
3301 College Ave
Fort Lauderdale, FL 33314

For contractual/administrative matters:
Northwestern University
Office for Sponsored Research
750 N. Lake Shore Dr., 7th Floor
Chicago, IL 60611
Attn: Manny S. Robert

7. AMENDMENT.

This Agreement represents the entire understanding of the parties with respect to the subject matter hereof. This Agreement may only be extended, renewed or otherwise amended by the mutual written consent of duly authorized representatives the parties hereto.

8. TERMINATION.

This Agreement shall be effective on the Effective Date and shall continue until the Agreement is terminated in accordance with the provisions hereof. Provider may terminate this Agreement at any time upon thirty (30) days' prior written notice to Recipient. This Agreement shall terminate automatically on

the date that is thirty (30) days after the Study has been completed. Upon Provider's knowledge of a material breach by Recipient, Provider may, in its sole discretion, either (1) provide Recipient with notice of and an opportunity to cure such breach and then terminate this Agreement if Recipient does not cure the breach within time period specified by Provider, or (2) terminate this Agreement immediately. Upon termination or expiration of this Agreement, Recipient shall cease use of all Data, and Recipient shall return or destroy all Data. Recipient may terminate this Agreement by returning or destroying the Data and providing written notice to Provider. If return or destruction of the Data is not feasible, Recipient shall continue the protections required under this Agreement for the Data consistent with the requirements of this Agreement and applicable HIPAA privacy standards.

9. MISCELLANEOUS.

a. Neither party shall use the other party's name, trademarks, or other logos in any publicity, advertising, or news release without the prior written approval of an authorized representative of that party. The parties agree that each party may disclose factual information regarding the existence and purpose of the relationship that is the subject of this Agreement for other purposes without written permission from the other party provided that any such statement shall accurately and appropriately describe the relationship of the parties and shall not in any manner imply endorsement by the other party whose name is being used.

b. The invalidity or unenforceability of any term or provision of this Agreement shall not affect the validity or enforceability of any other term or provision hereof.

c. Neither this Agreement nor the rights or obligations hereunder shall be assignable or otherwise transferred or subcontracted without the other party's prior written consent.

IN WITNESS WHEREOF, the Parties have caused this Agreement to be duly authorized and executed as of the Effective Date.

RECIPIENT

By: _____

Name: Gary Margules
(Authorized Official)

Title: VP for Office of Research and Technology Transfer

Date: 7/10/19

Read and Understood:

By: _____
Brad Callan, Investigator

Date: 7/10/19

PROVIDER

By: _____

Name: Manny S. Robert
(Authorized Official)

Title: Sr. Contract and Grant Officer

Date: _____

Read and Understood:

By: _____
James M. Elliott, PT, PhD

Date: _____

Appendix C: Demographics Collection Form

INITIAL QUESTIONNAIRE

Please answer all questions.

PERSONAL DETAILS

Title:	Name:	D.O.B.:	Gender: M / F
Address:			
		Postcode:	
Phone (H):		Phone (W):	
Mobile/Cell:		Email:	
Social-Security-Number (SSN): *			

**** The Accounting Services at Northwestern University will be given your name, address, and Social Security Number in order to issue a check for your study participation. Study payments are considered taxable income and reportable to the IRS. A Form 1099 will be sent to you if your total payments are $600 or more in a calendar year.***

Please fill in the contact details for 2 people (a relative, and a friend) who are not living with you. This information is only to be used to help us to find you if we lose contact.

Name:	Relationship:
Address:	
	Postcode:
Phone (H):	Phone (W):
Mobile:	Email:

Name:	Relationship:
Address:	
	Postcode:
Phone (H):	Phone (W):

Mobile:	Email:

Please fill in the contact details for your regular Primary Care Physician (PCP) below

PCP Name:	Phone
Address:	
	Postcode:

ACCIDENT HISTORY

Date of Accident: _____/_____/_____

1. At the time of the accident were you:

❑ The driver ❑ The front seat passenger

❑ A back seat passenger ❑ Riding a motorbike

2. Did you know the accident was coming? ❑ Yes ❑ No

3. Was the collision: ❑ Rear end ❑ Rear and front end ❑ Front end ❑ Side impact

4. Was the vehicle you were in stationary at the time of impact? ❑ Yes ❑ No

INITIAL SYMPTOMS

1. Following the car accident, did your neck pain start: ❑ Immediately ❑ Within 24 hours ❑ After 24 hours

2. Was your neck movement restricted following the accident? ❑ Not at all ❑ Mildly ❑ Moderately ❑ Severely

3. Did the restriction in movement start: ❑ Immediately ❑ Within 24 hours

❑ After 24 hours

4. **Did you lose consciousness immediately after the accident?** ❑ Yes ❑ No

5. **Were you admitted to hospital after the accident?** ❑ Yes ❑ No

MEDICAL HISTORY

1. **Have you had any major surgery or other injuries (e.g. back or neck pain)?** ❑ Yes ❑ No

*If **YES** – Please give details:*_____

_____

_____

_____

2. **Do you have any other medical conditions?** ❑ Yes ❑ No

*If **YES** – Please give details:*_____

_____

_____

_____

3. **Have any of the following investigations been performed for your neck pain?**

❑ Yes ❑ No

❑ X-ray – Please give
results:_____

_____

_____

❑ CT – Please give
results:_____

_____

_____

❑ MRI – Please give
results:_____

_____

_____

4. What treatments have you received for your neck pain since your accident?

Type of treatment	**Number of sessions**
❑ Physical Therapy	_____
❑ Chiropractic	_____
❑ Massage	_____
❑ Acupuncture	_____
❑ Other: _____	_____
❑ Other: _____	_____

❑ Surgical procedures

Provide details:

_____

_____

_____

_____

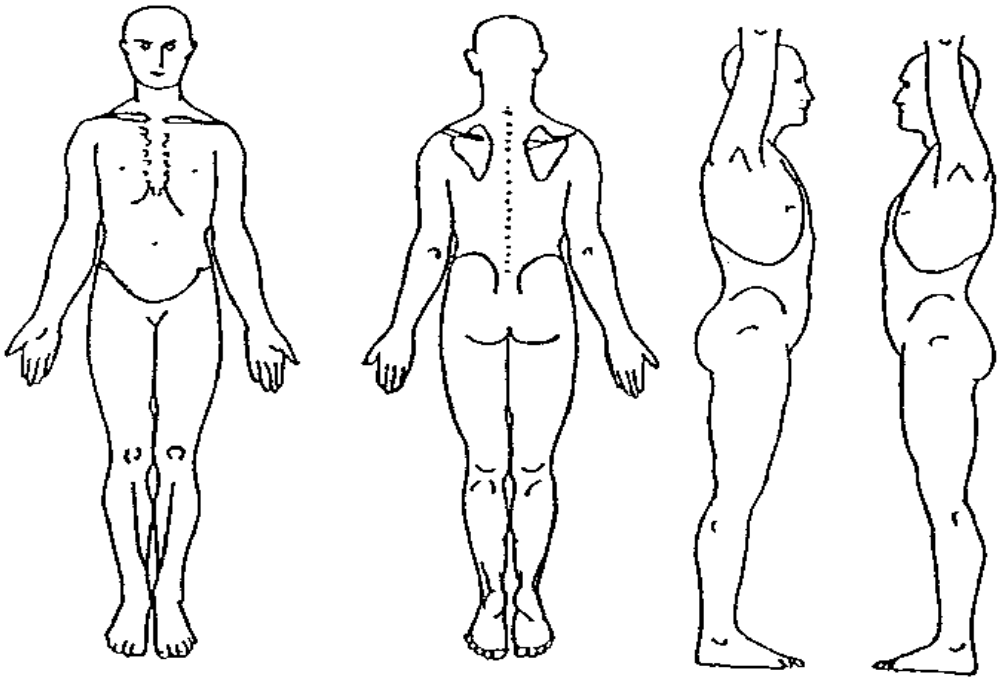
This question can assist us in understanding how inflammation can affect your condition.

5. Approximate date of your last normal menstrual period *(females only):*
_____/_____/_____

Check if not applicable: ❑

SYMPTOMS

1. Please mark on the body chart below where you feel pain or any other symptoms e.g. pins and needles or numbness.



2. On the scale below please estimate the typical intensity of your neck pain over the past 24 hours.

No Pain ——————————————— Worst Pain Imaginable

3. With respect to your whiplash injury, compared to straight after you had your accident, how would you describe yourself these days? (circle the most appropriate)

-5 -4 -3 -2 -1 0 1 2 3 4 5

vastly worse — unchanged — completely recovered

4. Headaches

i. Do you usually have headaches (i.e. prior to the accident)? ❑ Yes ❑ No

*If **YES**, since the whiplash injury, are the usual headaches:* ❑ The same

❑ Worse

ii. Do you have whiplash related headaches? ❑ Yes ❑ No

5. Dizziness

i. Do you usually experience dizziness / unsteadiness (i.e. prior to the accident)?

❑ Yes

❑ No

*If **YES**, is the dizziness / unsteadiness since the whiplash injury?*

❑ The same

❑ Worse

ii. Do you have whiplash related dizziness / unsteadiness? ❑ Yes ❑ No

Please be assured that the information asked on this form will remain anonymous at all times. It is collected for statistical purposes only.

1. Working Status:

❑ Employed (Occupation _____)

❑ Self Employed (Occupation _____)

❑ Home Duties

❑ Unemployed

❑ Retired

2. If employed/self employed, are you:

❑ Currently working usual hours

❑ Working reduced hours due to whiplash injury

IF REDUCED, indicate percentage of usual hours _____%

❑ Not working due to whiplash injury

3. Have you lodged a compensation claim:

❑ No
❑ Worker's compensation

❑ Third Party claim

❑ Other _____

4. Have you engaged the services of a solicitor/lawyer? ❑ Yes ❑ No

5. Education:

Please indicate the highest grade you completed at school: _____ _____

Please list any further qualifications and/or trades you have gained since leaving school, including their full-time duration:

Qualification 1: _____ Length of Course _____

Qualification 2: _____ Length of Course _____

Qualification 3: _____ Length of Course _____

Qualification 4: _____ Length of Course _____

MEDICATIONS

List medications you are taking for:

Name *Strength* *Number*

1. The whiplash injury

2. Other medical conditions

Appendix D: Informed Consent

Northwestern University
Department of Physical Therapy & Human Movement Sciences
Consent Form and HIPAA Authorization for Research
PROTOCOL TITLE: Neuromuscular Mechanisms Underlying Poor Recovery from Whiplash Injuries
PRINCIPAL INVESTIGATOR: James M. Elliott, PT, PhD
SUPPORTED BY: N National Institutes of Health (NIH); R01HD079076

Introduction

You are being asked to take part in a research study. This document has important information about the reason for the study, what you will do if you choose to be in this research study, and the way we would like to use information about you and your health.

Conflict of Interest Disclosure

The following disclosure is made to give you an opportunity to decide if this relationship will affect your willingness to participate in this research study:
Dr. Elliott, the person responsible for the conduct of this research study and another study member, Professor Todd Parrish, have equity interests in a medical consulting company, PainID, LLC. You have the right to discuss this study with another person who is not part of the research team before making your decision whether or not to be in the study.

What is the reason for doing this study?

The primary purpose of this study is to increase our understanding of injury to the spinal cord, and how this is related to muscle weakness and changes in muscles, reflexes and other symptoms you may have. We are also interested in researching the molecular changes within a blood sample that are associated with the body's response to stress and the recovery process after a motor vehicle collision (MVC) It is hoped that such knowledge will help us develop more effective treatment methods for spinal cord injury and other trauma to the spine, such as whiplash. You are being asked to participate in this research study because you had a MVC that resulted in neck pain.

How many people will take part in this study?

The study investigators hope to enroll 100 subjects at Northwestern University.
IRB #: STU00090769-MODCR0003 Approved by NU IRB for use on or after 10/19/2017 through 10/18/2018.
5/15/17

What will you do if you choose to be in this study?

Screening:

If you agree to participate, you will undergo a series of questions performed by one of the study's investigators to determine if you are eligible to participate in this study. The screening will last approximately 10 minutes and requires you to answer a series of questions related to your current pain levels and provide information related to your MVC. Also, pertinent medical history will be discussed. If your participation is not appropriate or safe, you cannot participate in this study.

Testing:

This study will involve two types of testing and one blood draw. The MRI portion of the study will take place at the Northwestern University Department of Radiology (710 N. Fairbanks Ct. Chicago IL. 60611 suite LC) and will involve imaging of your neck and legs. Each testing

session will last approximately 1 hour and will take place at less than one-week post MVC and then again at 2-weeks, 3-months and 1-year post MVC. The second portion of the study will take place at 645 N Michigan Ave, 8th floor. This testing may be performed on the same or on separate days, and involves testing of your reflexes/muscle strength. The testing session will last approximately 1 hour and will take place at 2-weeks, 3-months, and 1-year post MVC. The two types of sessions should be completed within 1 week of each other. Your total participation will involve a total of four MRI sessions and three reflex testing sessions over a time period of a year.
This study uses magnetic resonance imaging (MRI) to look at the spinal cord and muscles throughout the body. Magnetic resonance imaging is a type of imaging scan that use magnetic fields and radio waves to make a picture of the spinal cord and surrounding muscles allowing us to look the structure of the muscles throughout your body.
In order to make sure the MRI procedures will be safe, you will be asked to fill out a screening form before starting the study. It is important that you tell the researchers in this study if you have any history of:
 Metal fragments in your eyes or face.
 Implantation of any electronic devices such as (but not limited to) cardiac pacemakers, cardiac defibrillators, cochlear implants or nerve stimulators.
 Surgery on the blood vessels of your brain or the valves of the heart
 Claustrophobia (fear of enclosed places)
 Pregnant
The following is a more specific description of what is involved in this study:
Two small blood samples will be collected from you using a single blood draw, for a total of one tablespoon of blood. These samples can be obtained at the time blood is drawn by the ED staff as part of your clinical evaluation, or as a separate draw. If a separate blood draw, this will be done by either the ED tech or by the research associate who is trained in phlebotomy. Please note the blood draw is an optional element of the study. You may still participate in the study if you refuse to have your blood drawn.
IRB #: STU00090769-MODCR0003 Approved by NU IRB for use on or after 10/19/2017 through 10/18/2018.
5/15/17
Research with blood, tissue or body fluids (specimens) can help researchers understand how the human body works. Research can also answer other questions by using specimens. Researchers may develop new tests to find diseases, or new ways to treat diseases. In the future, research may help to develop new products, such as drugs. Specimens are commonly used for genetic research. Sometimes researchers collect and store many specimens together and use them for different kinds of research, or share them with other scientists; this is called a specimen repository or “biobank.”
The purpose of this particular repository or biobank is to have the opportunity in the future to look for genes, the expression of genes, proteins, or metabolites, or other molecular differences within blood samples that are associated with the body’s response to stress and the recovery process after motor vehicle collision. The specimens will be coded with a number, and will not identify you. The specimen number may be linked to certain information about you (such as age and sex), but will not be linked to personally identifying information such as your name, address, or social security number. It is not known at this time who (researchers) may be allowed access to the samples. In the future, the above biological data may be shared with commercial entities if this would allow the results of this research to be turned into useful products that would improve

recovery after motor vehicle collisions. If research done with your samples helps to develop new products in the future, you will not be paid. Because the samples will not identify you, you will be unable to withdraw them or have them destroyed.
These blood samples will be stored on the campus of Northwestern and later shipped in batches to a Biospecimen Processing Facility at the University of North Carolina. Samples will be banked indefinitely for future biological analyses examining the recovery process after motor vehicle collision.
In an effort to better understand the mechanical impacts of a motor vehicle accident and injuries to the neck, we are working with an accident reconstructionist. With your approval, we will ask for photos of the car accident and/or may send a member of the research team to take photos of the car. There will be no information on the photos that can be used to identify you. This is an optional element to the study. You may still participate if you do not agree to submit the photos.
You will be asked to undergo a MRI of your neck and leg muscles. We will ask you to complete a questionnaire related to the amount of pain and symptoms you are having at the time of the assessment as well as information related to your overall wellbeing. This questionnaire may be completed in a paper form or sent to you via a secure web application used for managing online surveys and databases. You will be asked to change into a hospital gown or surgical scrubs. Next, you will be asked to lie down on the scanner table. A moveable table will position you into the center of the magnet. In order to take the pictures, an imaging device will be placed over your neck and legs. You can speak to the researcher by talking out loud. If at any time or for any reason, you wish to stop the exam, you may do so by squeezing a rubber ball in your hand.
The MRI portion of the exam will take pictures of your spinal cord and the muscles in your leg and neck while you lay comfortably on your back inside the MRI tube. The images of your muscles allows us to observe any specific changes in the structure of your muscles as well as any changes that may occur inside your spinal cord; where the nerves that supply information to your muscles are situated. The expected time for this MRI scan is approximately one hour.
For the assessment of the strength in your leg muscles, you will be seated in an experimental chair with adjustable straps across your trunk and upper legs to comfortably maintain proper posture. Your foot will be comfortably secured into a rigid support that is connected to a
IRB #: STU00090769-MODCR0003 Approved by NU IRB for use on or after 10/19/2017 through 10/18/2018.
5/15/17
computer. The computer measures any muscle forces and torques that you create during a maximal effort contraction of your lower leg, similar to pushing your foot on the ‘gas pedal’ of your car and also lifting your toes up. Then, self-adhesive surface electrodes will be applied to your skin of your leg so that a brief electrical impulse can be applied while you produce a brief (4-5 seconds) maximal effort muscle contraction with your lower leg muscles. The expected time of this test is approximately 30 minutes.
For the assessment of reflexes, you will be seated in the same experimental chair as detailed above. Your foot will once again be comfortably secured into a rigid support that is connected to a computer. The computer measures any muscle forces and torques that you create with your leg muscles during a brief electrical impulse applied to your foot. Self-adhesive surface electrodes will be applied to your leg and used to measure any reflex muscle activity produced by the brief electrical stimulation applied to your foot. The expected time of this test is approximately 30 minutes.
The total expected time from your involvement in this study is two hours.

We will also ask that wear a small activity tracker that is clipped to your waistbandfor a period of two consecutive weeks. The device will record your step activity level. The monitoring device will be provided to you after the second and third visits. We ask that you mail back the pedometer to the study investigators via certified mail. You will be provided with the postage paid mailing materials at your second and third visits. Please note wearing the pedometer is an optional element of the study. You may still participate in the study if you decide not to wear the pedometer.

What are some of the possible risks and discomforts?

Your participation in this study may involve the following risks:

Venipuncture: The risks of taking blood include pain, a bruise at the point where the blood is taken,
redness and swelling of the vein and infection, and a rare risk of fainting. Care will be taken to avoid
these complications.

Surface electrodes: The self-adhesive surface electrodes used to record muscle activity may produce minor irritation of the skin. The possibility of irritation will be minimized by cleaning the skin with alcohol before and after application of the electrodes.

Electrical stimulation electrodes: The self-adhesive electrodes may induce some discomfort, although a comfortable level of stimulation will be found before performing the muscle contractions. There will be many rest periods to reduce the possibility of discomfort from the stimulation and the experiment will be discontinued if you report the significant discomfort.

Undergoing MRI scans may result in you feeling claustrophobic during the exam. In addition to feeling claustrophobic, some participants may find the experience to be loud in noise. However, our protocols include steps to ensure we significantly reduce the risk of these effects. For example, you will be provided with a hand-held rubber ball and instructed to squeeze should you feel uncomfortable during the exam. Once the ball is squeezed, an alarm will sound and the operator will stop the exam immediately. In addition, you will be provided with earplugs or earphones in order to reduce the noise of the MRI machine.

IRB #: STU00090769-MODCR0003 Approved by NU IRB for use on or after 10/19/2017 through 10/18/2018.
5/15/17

What are the Possible Benefits for Me or Others?

There may be no direct benefit to you by your participation in this research study. These procedures are entirely experimental and are not intended to provide any specific medical diagnosis or treatment. It is hoped that the results of this study will enhance our understanding and ultimately, treatments to prevent or reduce the development of chronic pain following whiplash injury.

What other procedures or courses of treatment might be available to me?

You have the alternative to choose not to participate in this research study. There are no satisfactory alternative procedures/assessments for this study. Of course, you may terminate your involvement in the experiment at any time for any reason.

Tissue or blood samples stored for future research

Allowing for the storage and future testing of your tissue and blood samples will involve no cost to you. Your sample will be used only for research and will not be sold. The research done with your tissue and blood sample may lead to the development of new products in the future. No compensation will be given to you now or in the future for the use of these samples.

Are there any financial costs to being in this study*?*
There are no costs to you for being in this study.
Will I receive payment for participation in this study?
You will receive $25 for completing each of the four MRI sessions. Another $25 will be paid for completing each of the three strength/reflex sessions. Once all sessions are completed, a total of $175 will be paid to you. Payments will be made by cash or check. The check will be mailed via standard US mail approximately 3-4 weeks following each visit. If you decide to withdraw from the study you will be paid only for the testing sessions that you complete. The Accounting Services at Northwestern University will be given your name, address, and Social Security Number in order to issue a check for your study participation. Study payments are considered taxable income and reportable to the IRS. A Form 1099 will be sent to you if your total payments are $600 or more in a calendar year.
NOTE: In the event any clinically important abnormalities are discovered on the MRI, the PI will immediately request the images be reviewed by a board-certified radiologist who is not involved in the study. The results of which will be promptly sent to the participant who will be encouraged to follow-up with his/her primary care physician. The PI will cover any/all costs associated with the Radiologist interpretation of the films.
If I have questions or concerns about this research study, whom can I call?
IRB #: STU00090769-MODCR0003 Approved by NU IRB for use on or after 10/19/2017 through 10/18/2018.
5/15/17
You can call us with your questions or concerns. If you have any illness or injury during your time on this study, you should call us promptly. Dr. James Elliott is the person in charge of this research study. Dr. Elliott can be reached at 312-503-2304 during Monday through Friday 7am-5pm. You can also call the study coordinator, Marie Wasielewski, at 630-688-4193 with any questions about this research study.
What are my rights as a research subject?
If you choose to be in this study, you have the right to be treated with respect, including respect for your decision whether or not you wish to continue or stop being in the study. You are free to choose to stop being in the study at any time.
Choosing not to be in this study or to stop being in this study will not result in any penalty. Specifically, your choice not to be in this study will not negatively affect your right to any present or future medical treatment to which you are otherwise entitled.
If you want to speak with someone who is not directly involved in this research, or have questions about your rights as a research subject, please contact the Northwestern University Institutional Review Board (IRB) Office. You can call them at 312-503-9338.
You may also contact the Rehabilitation Institute of Chicago Corporate Compliance Office at 312-238-2805
What else do I need to know?
A federal law, called the Genetic Information Nondiscrimination Act (GINA), generally makes it illegal for health insurance companies, group health plans, and most employers to discriminate against you based on your genetic information. This law generally will protect you in the following ways:
 Health insurance companies and group health plans may not request your genetic information that we get from this research.
 Health insurance companies and group health plans may not use your genetic information

when making decisions regarding your eligibility or premiums.
 Employers with 15 or more employees may not use your genetic information that we get from this research when making a decision to hire, promote, or fire you or when setting the terms of your employment.
Be aware that this federal law does not protect you against genetic discrimination by companies that sell life insurance, disability insurance, or long-term care insurance.

HIPAA Authorization

We are committed to respect your privacy and to keep your personal information confidential. When choosing to take part in this study, you are giving us the permission to use your personal health information that includes health information in your medical records and information that can identify you. For example, personal health information may include your name, address, or phone number. Your health information we may collect and use for this research includes:
 All information in a medical record

 Medical history
 Lab tests, or certain health information indicating or relating to a particular condition as well diaries and questionnaires
 Genetic health information: DNA, RNA
Once we have the health information listed above, we may share some of this information with the following offices or entities outside of Northwestern University and its clinical partners (or affiliates): the Northwestern University Institutional Review Board Office and Office for Research Integrity; the US Office of Research Integrity; the US Office for Human Research Protections; the US Food and Drug Administration.
Any research information shared with outside entities will not contain your name, address, telephone or social security number or any other personal identifier unless disclosure of the identifier is necessary for review by such parties or is required by law or University policy [except that such information may be viewed by the Study sponsor and its partners or contractors at the Principal Investigator's office].
 Authorized members of the Northwestern University workforce, who may need to see your information, such as administrative staff members from the Office for Research, Office for Research Integrity and members of the Institutional Review Board.
 Clinical affiliates, including but not limited the Rehabilitation Institute of Chicago (RIC), Northwestern Medical Group (NMG), Northwestern Memorial Hospital (NMH), Northwestern Lake Forest Hospital (NLFH), and the Ann & Robert H. Lurie Children's Hospital of Chicago (Lurie Children's). Your participation in this clinical trial may be tracked in an electronic database and may be seen by investigators running other trials that you are enrolled in and by your healthcare providers.
 Other University research centers and University contractors who are also working on the study
Those persons who get your health information may not be required by Federal privacy laws (such as the Privacy Rule) to protect it. Some of those persons may be able to share your information with others without your separate permission.
The results of this study may also be used for teaching, publications, or for presentation at scientific meetings. A study code number rather than your name or other identifying information

will protect your identity.

Certificate of Confidentiality

The principal investigator has a Certificate of Confidentiality from the federal government. A Certificate of Confidentiality helps protect the privacy of human research participants enrolled in studies that collect sensitive information. Certificates protect against legal demands, such as court orders and subpoenas, for information that could identify you in this study. For additional information about Certificates of Confidentiality see http://grants.nih.gov/grants/policy/coc/faqs.htm.

We will do everything we can to keep others from learning about your participation in this study. To further help us protect your privacy, we have obtained a Certificate of Confidentiality from the United States Department of Health and Human Services (DHHS).

With this Certificate, we cannot be forced (for example by court order or subpoena) to disclose information that may identify you in any federal, state, local, civil, criminal, legislative, administrative, or other proceedings. The researchers will use the Certificate to resist any demands for information that would identify you, except to prevent serious harm to you or others, and as explained below.

You should understand that a Certificate of Confidentiality does not prevent you, or a member of your family, from voluntarily releasing information about yourself, or your involvement in this study.

If an insurer or employer learns about your participation, and obtains your consent to receive research information, then we may not use the Certificate of Confidentiality to withhold this information. This means that you and your family must also actively protect your own privacy.

You should understand that we will in all cases, take the necessary action, including reporting to authorities, to prevent serious harm to yourself, children, or others. For example, in the case of child abuse or neglect.

Please note that:

 You do not have to sign this consent form. If you do not, it will not affect your treatment by health care providers, or the payment or enrollment in any health plans, or affect your eligibility for benefits. However, you will not be allowed to take part in this research study.

 You may change your mind and "take back" (revoke) this consent at any time. Even if you revoke this consent, the Principal Investigator may still use or share health information that was obtained about you before you revoked your consent as needed for the purpose of this study. To revoke your consent for the use of your health information, you must do so in writing to:

James Elliott, PT, PhD
Northwestern University
Department of Physical Therapy and Human Movement Sciences
645 N Michigan Av Suite 1100
Chicago, IL 60611

 Unless you revoke your consent, it will not expire.

IRB #: STU00090769-MODCR0003 Approved by NU IRB for use on or after 10/19/2017 through 10/18/2018.
5/15/17

Optional Elements:
The following research activities are optional, meaning that you do not have to agree to them in order to participate in the research study. Please indicate your willingness to participate in these optional activities by placing your initials next to each activity.
I agree I disagree

_____ _____
The researcher may collect or take photos of the damaged vehicle to aid with data analysis. The researcher will not share these photos with anyone outside of the immediate study team.

_____ _____
The researcher may take a blood sample taken during the study. These samples will be retained in non-identifiable form, meaning that there will be no information associated with the blood or samples that will allow anyone to readily ascertain my identity.

_____ _____
The researcher may ask me to wear a pedometer clipped to my waistband for two weeks to record my physical activity. The information collected will not have any personal identifying information and will not be shared with anyone outside the immediate study team.

_____ _____
The researcher may contact me in the future to see whether I am interested in participating in other research studies.
IRB #: STU00090769-MODCR0003 Approved by NU IRB for use on or after 10/19/2017 through 10/18/2018.
5/15/17

Appendix E: Neck Disability Index

Neck Disability Index

This questionnaire has been designed to give us information as to how your neck pain has affected your ability to manage in everyday life. Please answer every section and **mark in each section only the one box that applies to you**. We realise you may consider that two or more statements in any one section relate to you, but please just mark the box that most closely describes your problem.

Section 1: Pain Intensity

£ I have no pain at the moment
£ The pain is very mild at the moment
£ The pain is moderate at the moment
£ The pain is fairly severe at the moment
£ The pain is very severe at the moment
£ The pain is the worst imaginable at the moment

Section 2: Personal Care (Washing, Dressing, etc.)

£ I can look after myself normally without causing extra pain
£ I can look after myself normally but it causes extra pain
£ It is painful to look after myself and I am slow and careful
£ I need some help but can manage most of my personal care
£ I need help every day in most aspects of self care
£ I do not get dressed, I wash with difficulty and stay in bed

Section 3: Lifting

£ I can lift heavy weights without extra pain
£ I can lift heavy weights but it gives extra pain
£ Pain prevents me lifting heavy weights off the floor, but I can manage if they are conveniently placed, for example on a table
£ Pain prevents me from lifting heavy weights but I can manage light to medium weights if they are conveniently positioned
£ I can only lift very light weights
£ I cannot lift or carry anything

Section 4: Reading

£ I can read as much as I want to with no pain in my neck
£ I can read as much as I want to with slight pain in my neck
£ I can read as much as I want with moderate pain in my neck
£ I can't read as much as I want because of moderate pain in my neck
£ I can hardly read at all because of severe pain in my neck
£ I cannot read at all

Section 5: Headaches

£ I have no headaches at all
£ I have slight headaches, which come infrequently
£ I have moderate headaches, which come infrequently
£ I have moderate headaches, which come frequently
£ I have severe headaches, which come frequently
£ I have headaches almost all the time

Section 6: Concentration

£ I can concentrate fully when I want to with no difficulty
£ I can concentrate fully when I want to with slight difficulty
£ I have a fair degree of difficulty in concentrating when I want to
£ I have a lot of difficulty in concentrating when I want to
£ I have a great deal of difficulty in concentrating when I want to
£ I cannot concentrate at all

Section 7: Work
£ I can do as much work as I want to
£ I can only do my usual work, but no more
£ I can do most of my usual work, but no more
£ I cannot do my usual work
£ I can hardly do any work at all
£ I can't do any work at all
Section 8: Driving
£ I can drive my car without any neck pain
£ I can drive my car as long as I want with slight pain in my neck
£ I can drive my car as long as I want with moderate pain in my neck
£ I can't drive my car as long as I want because of moderate pain in my neck
£ I can hardly drive at all because of severe pain in my neck
£ I can't drive my car at all
Section 9: Sleeping
£ I have no trouble sleeping
£ My sleep is slightly disturbed (less than 1 hr sleepless)
£ My sleep is mildly disturbed (1-2 hrs sleepless)
£ My sleep is moderately disturbed (2-3 hrs sleepless)
£ My sleep is greatly disturbed (3-5 hrs sleepless)
£ My sleep is completely disturbed (5-7 hrs sleepless)
Section 10: Recreation
£ I am able to engage in all my recreation activities with no neck pain at all
£ I am able to engage in all my recreation activities, with some pain in my neck
£ I am able to engage in most, but not all of my usual recreation activities because of pain in my neck
£ I am able to engage in a few of my usual recreation activities because of pain in my neck
£ I can hardly do any recreation activities because of pain in my neck
£ I can't do any recreation activities at all
Score: /50 Transform to percentage score x 100 = %points
Scoring: For each section the total possible score is 5: if the first statement is marked the section score = 0, if the last statement is marked it = 5. If all ten sections are completed the score is calculated as follows: Example:16 (total scored)
50 (total possible score) x 100 = 32%
If one section is missed or not applicable the score is calculated: 16 (total scored)
45 (total possible score) x 100 = 35.5%
Minimum Detectable Change (90% confidence): 5 points or 10 %points
NDI developed by: Vernon, H. & Mior, S. (1991). The Neck Disability Index: A study of reliability and validity. Journal of Manipulative and Physiological Therapeutics. 14, 409-415