10-15-2015 12:00 AM

Inter-rater Reliability of the McKenzie System of Mechanical Diagnosis and Therapy in the Examination of the Knee

Sean Willis, *The University of Western Ontario*

Supervisor: Dr. Shawn Robbins, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Health and Rehabilitation Sciences

INTER-RATER RELIABILITY OF THE MCKENZIE SYSTEM OF MECHANICAL DIAGNOSIS AND THERAPY IN THE EXAMINATION OF THE KNEE

(Thesis format: Monograph)

by

Sean Willis

Graduate Program in Health and Rehabilitation Sciences

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
Health and Rehabilitation Sciences (Physical Therapy)

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

Objective: The purpose of this thesis was to investigate the inter-rater reliability of the McKenzie System of Mechanical Diagnosis and Therapy (MDT) when classifying patients with musculoskeletal knee pain using clinical vignettes. **Methods:** This study was divided into two phases. First, ten clinicians experienced in the use of MDT were randomly recruited to write a total of 60 clinical vignettes based upon the initial assessment of past patients with knee pain. Second, six different MDT raters were recruited to rate 53 selected vignettes and reliability was determined using Fleiss Kappa. **Results:** There was "substantial agreement" among six MDT raters classifying the clinical vignettes into one of four categories (kappa=0.72). There was no statistically significant difference between therapists with different levels of training. **Significance:** These findings indicate that the McKenzie System of MDT is a reliable method of classifying patients presenting with musculoskeletal knee pain when using clinical vignettes.

Keywords

McKenzie system, knee, Mechanical Diagnosis and Therapy, reliability, vignette

Co-Authorship Statement

Each of the four chapters in this thesis is based upon research designed and organized by the author. Members of the advisory committee provided regular guidance and feedback during this process. All chapters were authored primarily by Sean Willis with editing provided by the advisory committee.

Acknowledgments

I would like to sincerely thank to my supervisors, Dr. Shawn Robbins and Dr. Trevor Birmingham for their patience, guidance and dedication throughout my Masters degree. Their direction and encouragement were essential for the successful completion of this project.

My sincere gratitude is extended to the members of my advisor committee, Richard Rosedale and Ravi Rastogi. Your leadership, expertise, ideas and encouragement were truly invaluable. There are not enough words to convey my appreciation and thanks.

I would also like to thank members of the MDT community that agreed to participate. I know as a busy clinician how difficult it can be to take on work above your patient care. I extend my sincerest appreciation to you and your contributions to making this project possible.

Finally, I would like to acknowledge my wife, Darlene, and my boys, Benjamin and Andrew. Without your love and support, success in this endeavor would not have been possible. I love you and thank you so very much for being there when I needed you the most.

Table of Contents

Abstract ii

Co-Authorship Statement iii

Acknowledgments iv

Table of Contents v

List of Tables vii

List of Appendices viii

List of Abbreviations ix

Chapter 1 1

1 Overview of Problem 1

1.1 Purpose 3

1.2 Structure of Thesis 3

Chapter 2 4

2 Literature Review 4

2.1 Medical Imaging and Asymptomatic Pathology 4

2.2 Orthopaedic Special Tests (OSTs) in the Examination of the Knee 5

2.3 McKenzie System of Mechanical Diagnosis and Therapy 8

2.4 Clinical Vignettes 17

2.5 Research Gaps 19

Chapter 3 20

3 Inter-rater Reliability of the McKenzie System of Mechanical Diagnosis and Therapy in the Examination of the Knee 20

3.1 Study Objectives 20

3.2 General Study Design 20

3.2.1 Sample Size 21

3.2.2 Phase 1 21
3.2.3 Phase 2 23
3.3 Analysis 24
3.4 Results 25
Chapter 4 30
4 Discussion 30
4.1 Overview 30
4.2 Key Findings of the Thesis Project 30
4.3 Clinical Implications 32
4.4 Limitations 33
4.5 Future Recommendations 34
4.6 Conclusion 35
References 36
Appendices 41

List of Tables

Table 1 Summary of MDT classifications and clinical presentation 13

Table 2 Subgroups of MDT OTHER classification 14

Table 3 Demographic information of participating phase 2 raters (n=6) 26

Table 4 Frequency distribution of category of classification by individual rater 28

Table 5 Percentage agreement and kappa (standard error) for paired comparisons among the six raters 28

Table 6 Percentage agreement and kappa (standard error) of individual raters versus the provisional classification grouped by MDT education 29

Table 7 Reliability by MDT education across raters 29

List of Appendices

Appendix A Ethics Approval from Western University Health Science Research Ethics Board 41

Appendix B Phase 1 Recruitment Email 42

Appendix C Phase 1 Letter of Information and Consent Form 43

Appendix D McKenzie Institute Lower Extremities Assessment Form 49

Appendix E Derangement Vignette 51

Appendix F Articular Dysfunction Vignette 53

Appendix G Contractile Dysfunction Vignette 55

Appendix H OTHER Vignette 57

Appendix I Phase 2 Recruitment Email 59

Appendix J Phase 2 Letter of Information and Consent Form 60

Appendix K Phase 2 Data Collection Form 66

List of Abbreviations

ACL – anterior cruciate ligament

CI – confidence interval

LR – likelihood ratio

MDT – McKenzie System of Mechanical Diagnosis and Therapy

MRI – magnetic resonance imaging

NBA – National Basketball Association

OA – osteoarthritis

OSTs – orthopaedic special tests

PFPS – Patellofemoral Pain Syndrome

QUADAS – Quality Assessment of Diagnostic Accuracy Studies

SD – standard deviation

SE – standard error

Chapter 1

1 Overview of Problem

In the US, the prevalence of knee pain has increased by 65% over the last 20 years (Nguyen et al., 2011). Over 4 billion dollars are spent annually on arthroscopic knee surgery alone (Gage, McIlvain, Collins, Fields & Comstock, 2012) despite evidence suggesting arthroscopic surgery does not result in superior patient outcomes (Kirkley et al, 2008; Sihovnen et al, 2013; Thorlund, Juhl, Roos & Lohmander, 2015). Osteoarthritis (OA) related knee pain has been identified as a possible trigger for physical and functional decline for older adults (Jinks, Jordan & Croft, 2007). For those suffering from OA related hip and knee pain the cost of time lost from employment and leisure as well as their unpaid caregivers is often underestimated in the contribution to the overall burden of OA (Gupta, Hawker, Laporte, Croxford & Coyte, 2005). Within the current environment of fiscal responsibility in healthcare, it is vital that the overall costs of knee pain are recognized so that interventions that reduce the physical and financial burden are identified and funded to maximize patient outcomes. For this to occur, clinicians must possess the skills or use methods of assessment that have the clinical utility to identify the most appropriate, cost effective intervention from which the patient will benefit. Therefore, it is essential that an orthopaedic evaluation of the knee is valid and reliable, is guided by clear diagnostic criteria and provides the clinician with prognostic value.

An established body of evidence highlights limitations with the diagnostic validity of orthopaedic special tests (OSTs) used in the clinical examination of the knee (Cook, Mabry, Reiman & Hegedus, 2012; Geraets et al., 2015; Hegedus, Cook, Hasselblad, Goode & McRory, 2007; Lange et al., 2014; Leblanc et al., 2015; Peeler, Leiter, & MacDonald, 2010). The reported psychometric properties and hence the diagnostic accuracy of many of the commonly used OSTs are influenced by a number of factors including but not limited to rater experience, varied interpretation of the result findings, lack of a standardized approach to performance of the test and study design-related bias (Cook et al., 2012; Hegedus et al., 2007; Geraets et al., 2015; Lange et al., 2014; Leblanc et al., 2015; Peeler et al., 2010). Moreover, research around medical imaging such as

magnetic resonance imaging (MRI) have investigated the limitation of findings through the presence of asymptomatic pathology and abnormalities (Beattie et al., 2008; Boks, Vroegindeweij, Koes, Hunink &Bierma-Zeinstra, 2006; Kaplan, Schurhoff, Selesnick, Thorpe & Uribe, 2005; LaPrade, Burnett, Veenstra & Hodgman, 1994). With the relationship between pain and radiographic pathology not fully understood and limitations in reported diagnostic accuracy of OSTs, it has been suggested that specifically defined criteria used by non-pathoanatomical classification systems may offer better utility and should be considered as an alternative to the current model (Rosedale et al, 2014).

One system that has not been thoroughly tested for use with musculoskeletal pain in the extremity is the McKenzie System of Mechanical Diagnosis and Therapy (MDT). The MDT system of classification uses a non-pathoanatomically specific approach to classify patients based on their response to repeated end range loading strategies. Although demonstrating good inter-rater reliability in the assessment of musculoskeletal spinal pain (Clare, Adams & Maher, 2005; Kilpikoski, Airaksinen, Kankaanpaa, Leminen, Videman & Alen, 2002; Razmjou, Kramer & Yamada, 2000), shoulder pain (Heider Abady, Rosedale, Overend, Chesworth & Rotondi, 2014) and the extremities (Kelly, May, & Ross, 2008; May & Ross, 2009), MDT has not been evaluated on its use in the assessment of musculoskeletal knee pain.

Clinical vignette based methodologies are often used in the evaluation of decision making and clinical judgment of health professionals (Evans et al., 2015). Although often criticized because they do not reflect actual practice which may influence results and conclusions of studies, well designed vignette studies can be practical, offer flexibility, avoid ethical and observational issues and be generalizable to real world settings (Evans et al., 2015; Peabody, Luck, Glassman, Dresselhaus & Lee, 2000; Rutten, Harting, Rutten, Bekkering & Kremers, 2006). Clinical vignettes are an inexpensive option to control multiple variables, collect information simultaneously from multiple sources, and isolate clinical decision making. Thus they can provide an initial step in the investigation of the reliability of MDT for knee conditions.

1.1 Purpose

Considering the classification system of MDT has not been rigorously tested in the assessment of musculoskeletal knee pain and the use of clinical vignette based methodologies are a valid approach to examine clinical decision making, the purpose of this study is to determine the inter-rater reliability of MDT in the examination of the knee and what influence the level of MDT training may have on reliability.

1.2 Structure of Thesis

This document is presented in the "monograph" format described by the Western University Faculty of Graduate Studies.

In Chapter 2, a review of the literature is performed examining medical imaging and the presence of asymptomatic pathology, OSTs in the examination of the knee, the validity and reliability of MDT and the use of clinical vignettes in medical research.

Chapter 3 outlines the methods and presents results of the study. Chapter 4 discusses study findings, implications for health care professionals and limitations of the study. Recommendations for future research in this area are outlined.

Chapter 2

2 Literature Review

This chapter reviews the key findings and conclusions from a literature review in the areas of medical imaging and asymptomatic pathology, orthopaedic special tests (OSTs) in the examination of the knee, the McKenzie system of Mechanical Diagnosis and Therapy and the use of clinical vignettes in the research of clinical decision making. Gaps in the current research are also identified.

2.1 Medical Imaging and Asymptomatic Pathology

The diagnostic accuracy of an orthopaedic test is dependent on its ability to rule in or rule out pathology. The clinical utility of that test may be partially determined by the ability of that test to discriminate between symptomatic and asymptomatic pathology or abnormalities. Studying athletes and active individuals, a number of articles have highlighted the presence of previously undiagnosed anatomical abnormalities with medical imaging in pain-free individuals (Beattie et al., 2008; Boks et al., 2006; Kaplan et al., 2005; LaPrade et al., 1994).

Kaplan et al. (2005) reviewed the knee MRI findings of 20 National Basketball Association (NBA) players that met the inclusion criteria of no history of knee pain or surgery and had negative tests on physical examination for the presence of knee abnormalities such as meniscal and ligamentous disruptions and patella-femoral joint pain. The findings of the study, looking at players ranging from 21 to 36 years old, found that 47.5% of the evaluated knees had articular cartilage lesions and 20% of knees had meniscal tears. In their conclusion, the authors noted the influence of diagnostic imaging on clinical decisions and cautioned that findings do not indicate symptoms or functional level.

Investigating 100 patients with suspected meniscal tear, Zanetti, Pfirrmann, Schmid, Romero, Seifert and Hodler (2003) found 57 patients on MRI had a meniscal lesion on the symptomatic knee and of those, 36 had a meniscal lesion on the asymptomatic side

(63%). In a similar study, Boks et al. (2006) examined the MRI results of 134 patients with knee pain and found that of the 45 patients with a meniscal tear on the symptomatic side, 19 (42%) had one on the asymptomatic side. These values are substantially higher than those previously reported by LaPrade et al. (1994) who concluded that emphasis is needed on the importance of matching MRI findings with the history and physical examination after finding a prevalence rate of 5.6% for asymptomatic meniscal tear in 54 men and women with no previous history of knee pain or trauma. What should be noted is that Kaplan et al. (2005), LaPrade et al. (1994) and Zanetti et al. (2002), screened subjects for knee pathology prior to their participation.

Using a sample comprised of men and women of an average age of 41.5 years of age, Beattie et al. (2008), recruited subjects with no previous history or diagnosis of knee pathology to undergo a MRI and X-ray on their non-dominant knee. Although the prevalence of cartilage lesions was relatively low at 11%, all but one participant exhibited a meniscal abnormality in at least one region of the knee with more than 60% of participants having an abnormality in at least three of four regions.

In summary, it has been suggested that MRI findings should be interpreted with caution as findings do not indicate symptoms or functional level (Kaplan et al., 2005). With evidence demonstrating the presence of asymptomatic pathology in the knee, it is reasonable to question not only the diagnostic utility of detecting pathology but also whether one can be certain an implicated structure is the cause of an individual's symptoms. This can be of particular consequence when patients present with a history of pain and MRI identified pathology to which a decision on care must be made, often having to decide whether or not surgery is indicated.

2.2 Orthopaedic Special Tests (OSTs) in the Examination of the Knee

A change in practice has gradually occurred over the last several decades as clinicians performing an orthopaedic assessment have become over reliant on the results of OSTs and medical imaging (Cook, 2010). The psychometric properties and hence, the diagnostic accuracy of these OSTs is often influenced by a number of factors including

but not limited to rater experience, varied interpretation of the result findings and lack of a standardized approach to performance of the test (Cook et al., 2012; Hegedus et al., 2007; Peeler et al., 2010). Several threats to diagnostic validity such as selection bias, verification bias and the study sample have been identified and can inflate estimates of diagnostic accuracy (Lijmer et al., 1999). As a result, the usefulness of many of these tests has been questioned (Cook et al., 2012; Geraets et al., 2015; Goossens et al., 2015; Hegedus et al., 2007).

Investigating the diagnostic accuracy of three common ACL tests, Peeler et al. (2010) found only moderate levels of inter-rater agreement for the anterior drawer (0.57), the Lachman (0.45), and the pivot shift (0.53). The Lachman demonstrated a sensitivity of 83% with orthopaedic surgeons but varied greatly within clinician groups, family physicians and therapists, ranging from 15% to 87%. Peeler et al. concluded that variables such as level of experience and degree of training or specialization may impact the accuracy of testing. Geraets et al. (2015) had similar findings comparing an orthopaedic surgeon and primary care physician and the diagnostic value of the subjective and objective exams. They concluded that the objective exam, while improving an orthopaedic surgeons' positive predictive value of an ACL tear from 0.65 to 0.94, offered no value to the assessment for the primary care physician, dropping positive predictive value from 0.69 to 0.62.

Leblanc et al. (2015) have suggested that the clinical setting and the degree of tear will impact the diagnostic accuracy of tests for anterior knee instability. In their systematic review, they found that the sensitivity of the Lachman and pivot shift tests were lower when patients were awake versus under anesthetic and in the presence of a partial versus a complete tear. They also found insufficient data to calculate a pooled specificity and as a result, were unable to give a clear recommendation of the diagnostic accuracy of the physical examination in ACL deficient knees. Similarly, Lange et al. (2015) were unable to perform a meta-analysis during their systematic review of the physical tests for ACL rupture as a result of heterogeneity of the sample populations, the reliability measures used and the poor methodological quality of the studies reviewed.

Of the 18 studies qualifying to be included in a meta-analysis of the physical tests for meniscal tears, Hegedus et al. (2007) found three tests to be studied most: McMurray's, Apley's and joint line tenderness. Of those three tests, McMurray's demonstrated the highest sensitivity of 70% but also the lowest specificity at 71%. Joint line tenderness had the highest specificity with 77% but also the lowest sensitivity at 63%. Hegedus et al. concluded that no single test is able to accurately diagnosis a meniscal tear and recommended that the performance and interpretation of the tests be standardized.

More recently, Goossens et al. (2015) reported the Thessaly meniscal test with a sensitivity of 64% and a specificity of 53%. When combined with the McMurray's test, the sensitivity dropped to 53%. As a result, the authors concluded that either test in isolation or combined, does not appear useful in the detecting of meniscal tears. Further to this, it was recommended that research should focus on the development of a better diagnostic model of examination. Of interest, Campbell et al. (2014) investigated the correlation between location of preoperative knee pain and arthroscopic knee findings. The authors found that no significant correlation ($p=0.98$) existed between pain location and pathology and concluded that because of the varied nature of pain their results dispute the widely held beliefs that the location of pain is related to underlying pathology.

Cook et al. (2012) conducted a systematic review for clinical tests for screening and diagnosing patellofemoral pain syndrome (PFPS) using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) scoring for methodological quality. Of the 704 articles identified, 9 met selection criteria, presenting with 22 clinical tests for review. None of the 22 tests reviewed demonstrated a positive likelihood ratio (+LR) greater than 5.0 and a negative likelihood ratio (-LR) less than 0.20, meaning an inability to rule in or out (PFPS). Of those tests that had a stand-alone +LR greater than 5.0, those studies had the lowest methodological quality and/or used normals as the control group introducing quality bias and affecting diagnostic accuracy.

In summary, while widely accepted, the use of OSTs in the examination of the knee demonstrates limited ability to establish a clear relationship between clinical testing and symptomatic pathology. A potential explanation for this limitation may be the over-

reliance on identifying an anatomical structure or structures that are the cause of pain. For the ideal management of musculoskeletal problems, clinicians require the use of accurate tests and validated diagnostic criteria. Issues with the current approach highlight the need to explore other systems of clinical examination.

2.3 McKenzie System of Mechanical Diagnosis and Therapy

The McKenzie System of Mechanical Diagnosis and Therapy (MDT) is a non-pathoanatomically specific classification system that was originally developed for use in spinal conditions. The assessment screens out potential red flag issues such as fractures, neurological or vascular issues to determine if a patient's symptoms are mechanical in nature. MDT involves a detailed history and an examination in which baseline symptoms, both with function and at rest, are established and then re-evaluated following the patient performing repeated end range loading movements to the affected area. A key characteristic of the system that has shown potential as a prognostic indicator of musculoskeletal pain is directional preference (May & Aina, 2012). Directional preference is defined as the rapid improvement of a patient's symptoms with positioning or movement in one specific direction while commonly worsening with positioning or movement in the opposite direction (McKenzie and May, 2003). Based on the patients' response to the assessment and potential change in baseline symptoms, the clinician is able to formulate a provisional classification and provide directed treatment.

The system is based not on determining an anatomical diagnosis but rather classification into one of four categories, the first three being specific mechanical syndromes: Derangement, Dysfunction, Postural or OTHER. The mechanical syndromes were originally developed based on particular patterns of symptoms and responses to movement in the spine that were seen by the founder of MDT, Robin McKenzie. More recently, the MDT system has been used with increased frequency by trained clinicians in the evaluation of joints in the extremities.

The Derangement syndrome is the most common of the three mechanical syndromes and is varied in its clinical presentation; however, the key characteristic is the presence of a

directional preference with loading strategies (McKenzie and May, 2003). An example of this would be a worsening of a patient's symptoms with movements into flexion but an improvement or abolishment of symptoms with movements into extension. The clinical presentation for Dysfunction syndrome is intermittent pain that is consistently reproduced at the end-range of a restricted movement but will not persist once mechanical loading strategies have ceased (McKenzie and May, 2003). In the extremities, the Dysfunction syndrome can be disseminated further as Articular and Contractile Dysfunction. Contractile Dysfunction is characterized by pain brought on by active and resisted movements, where passive range of motion is generally preserved (McKenzie and May, 2000). Articular Dysfunction is distinguished from contractile through the loss of active and passive range of motion with pain being produced at the end of available range and absent during resisted testing (McKenzie and May, 2000). Postural syndrome is distinguished by local, intermittent pain without movement loss that is brought on by sustained postures and abolished with posture correction (McKenzie and May, 2003). An annual review of the MDT educational program is conducted by the International Education Committee of the McKenzie Institute International. Any revisions or changes to the definitions or criteria are brought about by updates to published research literature, through feedback from MDT Faculty and of the evaluation of the system from the Committee members. The most current summary of the classifications is provided in Table 1 and is presently in use in the MDT education manuals. The OTHER category is made up of 10 diagnostic subgroups which together complete the full MDT classification system where each subgroup has its own definition and diagnostic criteria (Table 2). Table 2 has been modified from the original publication by May and Rosedale (2012) to reflect the most recent revisions made by the International Education Committee.

The MDT system of education has two levels of clinical competence, Credentialed and Diploma. Credentialed clinicians have completed four post-graduate courses and successful passed a standardized written and practical examination. Having attained Credentialed status, clinicians can then go on to acquire Diploma which consists of one University semester theoretical component and 360 hours of clinical practice mentorship. Once completed, the clinician must then pass an oral examination to be awarded Diploma status (http://www.mckenzieinstitute.org/).

In the literature, a spinal assessment using MDT has been shown to have good inter-rater reliability (Clare et al., 2005; Kilpikoski et al. 2002; Razmjou et al., 2000). Using two MDT trained therapists, one Credentialed and one Diploma, Razmajou et al. (2000) investigated the interrater reliability of the MDT system during the assessment of 45 patients presenting with mechanical low back pain. They found the overall reliability between raters on mechanical syndrome classification to be substantial (kappa=0.70) with the Derangement classification to have the highest reliability of kappa=0.96. The agreement of syndrome classification was 93% between raters for all responses. These results are similar to those of Kilpikoski et al. (2002) that found an overall reliability between two MDT trained Diploma raters on the assessment of 39 patients to be moderate (kappa=0.6) with 95% agreement on syndrome classification. The majority of participants (90%) were classified into the Derangement syndrome.

Clare et al. (2005) examined the reliability of the MDT system in the classification of patients with lumbar and cervical pain between 14 raters, seven Credentialed and seven Diploma. The overall reliability for the classification of patients was substantial (kappa=0.84) with kappa=1.0 for lumbar patients and kappa=0.63 for cervical patients. Overall agreement amongst raters was high with 96% for the total patient pool. The majority of patients (66%) were again classified as Derangement syndrome.

More recently, Werneke et al. (2014) investigated the relationship between pre-Credentialed level of training and therapist agreement in the McKenzie lumbar classification. Forty-seven raters of various levels of pre-Credentialed MDT training assessed over 1600 patients and found an overall range of kappa=0.37 to 0.44 for classification into one of the mechanical syndromes despite an observed agreement of 86 to 91%. It has been suggested this paradox results from the sensitivity of the kappa statistic when the prevalence of a rating is either very high or very low and that the interpretation and reporting of the kappa statistic alone may result in conclusions that may be misleading (Bryt, Bishop & Carlin, 1993; Cicchetti & Feinstein, 1990; de Vet, Mokkink, Terwee, Hoekstra and Knol, 2013; Feinstein & Cicchetti, 1990). Indeed, for patients classified into one of the four mechanical syndrome classifications, Werneke et al. (2014) reported Derangement syndrome among examiners to range from 334 to 512

(81 to 86%) with the remaining three classifications ranging from 0 to 27 (0 to 4.6%). This skewed distribution elevates the probability of agreement due to chance alone and thereby lowers the value of the kappa statistic which represents the proportion of agreement greater than that expected by chance (O'Leary et al., 2014; Werneke et al., 2014)

Previously published studies investigating the value of MDT guided treatment in the extremity had been restricted to case studies documenting patients presenting with shoulder and knee pain (Aina & May, 2004; Littlewood & May, 2007; Lynch & May, 2013) and despite an increase in clinical use, research evaluating the MDT assessment for musculoskeletal extremity problems is limited. In a pilot study, Kelly, May & Ross (2008) examined the reliability of trained MDT clinicians classifying clinical vignettes based on patients with musculoskeletal disorders in the extremity. They found the agreement among three Credentialed raters for 11 vignettes to be kappa=0.70. These results are similar to May & Ross (2009) which investigated the reliability of the MDT assessment form for extremity conditions by using 25 clinical vignettes. They found an overall level of agreement to be 92% with a kappa of 0.83 among 97 Diploma trained therapists. There was little difference in reliability between upper (kappa=0.85) and lower extremity (kappa=0.80) cases.

Surveying Diploma therapists, May and Rosedale (2012) gathered data on the prevalence of mechanical syndromes and treatment strategies in use for the extremities. The most commonly used classifications for patients presenting with musculoskeletal knee pain were: Derangement (42.7%), Articular Dysfunction (3.9%), Contractile Dysfunction (8.7%) and OTHER (44.7%), 20% of which were post-surgery or post-trauma. Of interest, May and Rosedale (2012) found that 85.8% of initial classifications remained stable throughout the treatment episode. More recently, Heider Abady et al. (2014) demonstrated almost perfect reliability (kappa=0.90) between six Diploma raters when using MDT to classify 54 clinical vignettes of patients with musculoskeletal shoulder pain with an overall level of multi-rater agreement to be 96%. Of note, the highest level of agreement in this study was for Spinal, with the category of OTHER having the lowest level of agreement.

In a recent randomized control trial, Rosedale et al. (2014) sought to examine the effectiveness of exercise intervention determined through an MDT on patients diagnosed with end stage knee OA. Not only were patients readily classified as Derangement or not Derangement, but it could be inferred that the large effect size of d = 0.77 to 0.87 for all primary outcomes seen at two weeks by the intervention group is attributable to the classification and exercise matching determined by the MDT assessment. Although the results of this study are encouraging, the reliability of the MDT classification system had not yet been previously studied in the knee.

In summary, although the MDT system has been shown to be reliable for assessment of musculoskeletal pain in the spine and shoulder, no studies have been conducted on the reliability of the MDT system for musculoskeletal knee pain.

Table 1 Summary of MDT classifications and clinical presentation

MDT Classification	Clinical Presentation
Derangement	Varied in its clinical presentation; associated with mechanical obstruction of an affected joint; however, the key characteristic is the presence of a directional preference with loading strategies. Directional preference is defined as the rapid improvement of a patient's symptoms with positioning or movement in one specific direction while commonly worsening with positioning or movement in the opposite direction.
Dysfunction	Intermittent pain that is consistently reproduced at the end-range of a restricted movement but will not persist once mechanical loading strategies have ceased. In the extremities, the Dysfunction syndrome can be disseminated further as Articular and Contractile Dysfunction.
Articular Dysfunction	Distinguished from contractile through the loss of active and passive range of motion with pain being produced at the end of available range and absent during resisted testing.
Contractile Dysfunction	Characterized by pain brought on by active and resisted movements, where passive range of motion is generally preserved.
OTHER	Category is made up of 10 diagnostic subgroups which together complete the full MDT classification system where each subgroup has its own definition and diagnostic criteria. (Table 2)

Table 2 Subgroups of MDT OTHER classification

Serious Pathology (list not exhaustive)		
Category	**Clinical Findings (Red Flags)**	**Clinical Examples**
Cancer	Age >55, history of cancer, unexplained weight loss, progressive, not relieved by rest	Maybe primary site or metastases
Fracture	History of significant trauma (If osteoporosis present; minor trauma) Loss of function. All movements make symptoms worse.	
Infection	Fever, malaise, constant pain, all movements worsen	

Non-Serious Pathology Subgroups for OTHER classification			
Subgroup	**Definition**	**Criteria**	**Clinical Example**
Chronic Pain Syndrome	Pain-generating mechanism influenced by psychosocial factors or neurophysiological changes	Persistent widespread pain, aggravation with all activity, disproportionate pain response to mechanical stimuli, inappropriate beliefs and attitudes about pain.	Regional pain syndromes
Inflammatory	Inflammatory arthropathy	Constant pain, morning stiffness, excessive movements exacerbate symptoms	RA, sero-negative arthritis, some stages of OA

Subgroup	Definition	Criteria	Clinical Example
Mechanically Inconclusive	Unknown musculoskeletal pathology	Derangement, Dysfunction, Postural and subgroups of OTHER excluded. Symptoms affected by positions or movements BUT no recognizable pattern identified OR inconsistent symptomatic and mechanical responses on loading.	
Peripheral Nerve Entrapment	Peripheral nerve entrapment	No spinal symptoms. Local paraesthesia / anaesthesia. May have local muscle weakness.	Carpal tunnel syndrome, myalgia paraesthetica
Post-surgery	Presentation relates to recent surgery	Recent surgery and still in post-operative protocol period.	
Soft Tissue Disease Process	A fibroblastic or degenerative disease process affecting inert soft tissue with unknown or disputed aetiology	Each disease process has a unique clinical presentation, natural history and response to a variety of interventions.	Frozen shoulder, Dupuytren's, plantar fascia syndrome

Subgroup	Definition	Criteria	Clinical Example
Structurally Compromised	Soft tissue and/or bony changes compromising joint integrity	Mechanical symptoms (ROM restricted, clunking, locking, catching). May have sensation of instability. Long history of symptoms or history of trauma. Irreversible with conservative care.	Late stage OA, dislocation, labral tear, cruciate ligament rupture, irreducible meniscal tear
Trauma/Recovering Trauma	Recent trauma associated with onset of symptoms	Recent trauma associated with onset of constant symptoms / recent trauma associated with onset of symptoms, now improving and pain intermittent.	
Vascular	Symptoms induced by poor blood supply due to pressure increase in a closed anatomical space.	Below knee symptoms, predominantly in younger athletes. Consistently induced by exercise or activity. May have pain and /or paraesthesia in field of local cutaneous nerve and local swelling.	Compartment syndrome

Source: May, S. & Rosedale, R. A survey of the McKenzie classification system in the extremities: prevalence of mechanical syndromes and preferred loading strategies. Physical Therapy, 92(9), 1175-86.

2.4 Clinical Vignettes

To assess inter-rater agreement, measurement of a clinicians' performance must "ultimately rely on measures that are valid, reliable, inexpensive and manageable" (Rutten et al., 2006, p. 492). Two methods presently used in the literature to assess reliability include the use of real patients and clinical vignettes. Each method possesses its own strengths and weaknesses. Recruiting actual patients allows for subtle variability in patient presentation for similar musculoskeletal problems. Actual patients may permit a true expression of the nature of symptoms and responses to testing and potentially allow for better interpretation of the clinical interaction (May & Ross, 2009). Using patients to test inter-rater agreement may improve the realism and depth to the clinical scenario which may improve external validity and generalizability of the study and findings.

However, there are limitations to using real patients. For instance, real patients may make measurement by direct observation difficult to apply, especially in larger samples, can be expensive and time-consuming, and is potentially subject to a Hawthorne effect (Rutten et al., 2006). Use of real patients may result in insufficient case mix (Peabody et al., 2000) which may inadvertently introduce sampling bias, especially as it relates to MDT. The Derangement syndrome is the most common classification in the spine (78%, May, 2006) and in the extremity (37%, May & Rosedale, 2012). Because of the apparent prevalence of the Derangement syndrome, the random recruitment of patients may unintentionally create a homogenous sample that potentially would not include all relevant syndromes within the classification system.

Another option to evaluate inter-rater agreement is through the use of clinical vignettes. Clinical vignettes have a long history of use (Evans et al., 2015) and are defined as written patient case studies based on realistic scenarios where clinicians are given one or more questions asking what they may do if given the actual patient (Veloski, Tai, Evans & Nash, 2005). A number of studies have used vignettes as a primary method of data collection ranging from physical therapy adherence to guidelines (Rutten et al., 2006), best practices of physical and occupational therapist for young patients with cerebral palsy (Saleh et al., 2008), to measuring the quality of physician practice (Peabody et al., 2004). Using clinical vignettes for data collection provide the user the advantages of the

ability to simultaneously collect information from a number of subjects, manipulate multiple variables and create heterogeneous case mixing, avoid ethical issues, and avoid observer effects that can affect observational studies (Gould, 1996).

In a comparison of vignettes, standardized patients and chart abstraction, Peabody et al., (2000) used the three methods to evaluate physician competence and the quality of their practice. The authors concluded that clinical vignettes can be used in an outpatient setting to evaluate quality of care, may offer an inexpensive way to provide adequate case mix and can be a valid and comprehensive means to evaluate processes of care in clinical practice. These findings are consistent with studies by Dresselhaus, Peabody, Luck and Bertenthal (2004) and Veloski et al. (2005), who added that clinical vignettes are an effective way to isolate decision making. In a validation study, Peabody et al. (2004) found clinical vignettes to be a valid tool to provide case-mix variation, and "are particularly useful for comparing quality among and within sites and may be useful for longitudinal evaluations of interventions intended to change clinical practice" (p. 771).

Despite being seen as a valid measurement tool, there appears to be a lack of literature validating framework for the generation or creation of clinical vignettes. When appraising and evaluating articles that use clinical vignettes, Gould (1996) attempted to address this by ensuring certain features were present. Gould (1996) recommended that authors should address internal validity issues by developing vignettes based on existing literature and/or case study review, the scenarios should be tested to remove ambiguity and reviewed by an expert panel that possesses the knowledge and expertise to determine appropriateness of the vignette for the study. Atzmüller & Steiner (2010) proposed that researchers should generate more vignettes than needed and subsequently select those vignettes that would create the best sample with which to test.

Further to this, Veloski et al. (2005) suggested that the clinical scenarios follow the same natural flow of a clinical assessment, that the order of the information be logical and sequenced as though a clinician were performing it on an actual patient. The vignette should be written as such to minimize confusion, remain specific to the goal of testing the hypothetical situation and maintain a level of uncertainty that does preclude the

clinicians' ability to articulate their interpretation of the case. Well written, realistic vignettes should simulate aspects of real world scenarios, a facet of construct validity, offer enough variability which relates to the study's internal validity and produce results that are generalizable to real world situations, reflecting external validity (Evans et al., 2015).

2.5 Research Gaps

Although studies have been done to determine the reliability of the MDT system, none have investigated the reliability of the system in the knee. The use of real patients in a reliability study may result in an insufficient case mix which may inadvertently introduce sampling bias, especially as it relates to MDT. With the Derangement syndrome being the most common classification in the spine (78%, May, 2006) and in the extremity (37%, May & Rosedale, 2012), clinical vignettes would allow for the creation of a heterogeneous sample that would include all relevant syndromes within the classification system and thus avoid these issues.

Chapter 3

3 Inter-rater Reliability of the McKenzie System of Mechanical Diagnosis and Therapy in the Examination of the Knee

This chapter reviews the study objectives, design and the methods used to determine the inter-rater reliability of the McKenzie System of Mechanical Diagnosis and Therapy (MDT) in the examination of the knee. The results of the study are also reported in this chapter.

3.1 Study Objectives

The McKenzie System of MDT is a widely used method of classification and management of musculoskeletal problems. Although the McKenzie system has been investigated for its reliability and efficacy in the management of spinal pain, few studies have evaluated the system when applying it to musculoskeletal problems in the extremities, in particular the knee. The objectives of this study were to:

1. To develop 53 clinical vignettes of patients presenting with musculoskeletal knee pain using past patient data from the caseloads of 10 MDT Credentialed or Diploma clinicians which are based on the definitions of four clinical classifications.

2. To test the inter-rater reliability of six MDT-trained experienced clinicians when classifying patients with musculoskeletal knee pain into one of four MDT classifications using written clinical vignettes.

3. To investigate the influence of the level of MDT education on the reliability of classifying patients with musculoskeletal knee pain using written clinical vignettes.

3.2 General Study Design

To achieve these objectives, a two phase study was conducted. To achieve objective one, the first phase consisted of the recruitment of 10 MDT clinicians to develop 53 clinical vignettes representative of the prevalence of musculoskeletal knee pain classified through

MDT. To meet objectives two and three, the second phase required the recruitment of an additional six MDT raters to classify the patients represented in the clinical vignettes and measure the reliability and level of agreement among the MDT raters. Ethical approval for the study was obtained from the Health Science Research Ethics Board at Western University (Appendix A).

3.2.1 Sample Size

Rotondi and Donner (2012) proposed a method of calculating the sample size for studies measuring inter-rater agreement for multiple outcomes and raters. To arrive at an estimated sample size, kappa was set at 0.8 (0.7 lower limit, 0.9 upper limit) based on levels of agreement with previous work evaluating the reliability of MDT in the extremities (Heidar Abady et al., 2014; Kelly, May, & Ross, 2008; May & Ross, 2009). With an alpha of 0.05 for six raters (phase two) and using the prevalence of the four common syndromes of 0.4(Derangement), 0.4(OTHER), 0.1(Contractile Dysfunction), 0.1(Articular Dysfunction) as outlined by May and Rosedale (2012), a value of 53 was determined for the number of clinical vignettes required for phase two. The sample size was estimated using a program developed by Rotondi (2013) for the R Project for Statistical Computing.

3.2.2 Phase 1

3.2.2.1 Participants

For the first phase of the study, ten clinicians experienced in the use of MDT in the extremity were recruited based on previous willingness to participate in research. These clinicians were asked generate 60 clinical vignettes, six vignettes per clinician, classified into one of four classifications. The sample size of ten was chosen to minimize the burden of creating the vignettes on the consenting clinician. To be included, clinicians were Credentialed or Diplomat with the McKenzie institute with more than three years of experience applying MDT to musculoskeletal disorders of the extremity and be registered on the publicly available list of MDT practitioners practicing in the United States or Canada. Clinicians were excluded if they are unable to understand written and spoken English, unable to provide informed consent, or unable to follow the instructions for

generating the clinical vignettes. Correspondence was conducted and informed consent was obtained from each clinician through electronic mail (Appendix B – C). In total, 20 clinicians were approached for recruitment to which 10 consented for participation in this study.

3.2.2.2 Vignette Development

For clinical vignette development, Atzmüller & Steiner (2010), Evans et al. (2015), Gould (1996) and Veloski et al. (2005) have suggested that more vignettes should be generated than will be used, be reviewed by an expert panel to determine appropriateness and to select the best sample for testing, be written based on or relating to a case study or clinical experience and follow a similar structure and natural flow for all vignettes used. For this study, clinicians were asked to generate vignettes based on their past patient assessment files. Clinicians were instructed that each clinical vignette is to be de-identified to only include gender, age range (eg. 35 to 40 years old) and a category of occupation. The written vignettes would be characteristic of one of the four MDT classifications identified by May and Rosedale (2012) as most prominent in patients with musculoskeletal knee pain: Derangement, Articular Dysfunction, Contractile Dysfunction and OTHER. Clinicians were asked to write the clinical vignettes on a blank McKenzie extremity assessment form (Appendix D). The blank McKenzie extremity assessment form used was revised from the standard form to exclude entry areas for patient names and other identifying information. To correspond with established prevalence, each clinician was asked to submit 6 clinical vignettes consisting of two Derangements, one Articular Dysfunction, one Contractile Dysfunction and one OTHER. A summary of these classifications is provided in Table 1.

Once received, all vignettes were reviewed by the author and a member of the advisory committee (SW & RR). The first reviewer (SW) is a MDT Credentialed physiotherapist and has 17 years of clinical practice working with patients with musculoskeletal knee pain. The second reviewer (RR) is a MDT Diploma physiotherapist with 23 years of clinical experience and is Senior Faculty of the McKenzie Institute. The review of the cases was done to ensure that each vignette was complete, possessed characteristics of one of the four MDT classifications requested and that a level of ambiguity existed that

would be present in the situation of a clinical patient presentation. Any discrepancies were identified, flagged and discussed with the subject who developed the vignette to ensure and verify accuracy of the case. In some situations, clinicians were unable to submit one or more of the number of vignettes matching the requested classification(s) because of the lack of a past patient assessment(s) that represented that classification(s). In those instances, clinicians chose another past patient assessment with one of the other requested classification(s) to submit to fulfill their quota of 6 clinical vignettes. In total, 60 vignettes were received: 24 Derangement, 8 Articular Dysfunction, 8 Contractile Dysfunction and 20 OTHER. After a review of all vignettes was completed, 53 cases were selected that were representative of the established prevalence for use in phase two of the study. Of the 53 vignettes, 22 were Derangement, 7 Articular Dysfunction, 7 Contractile Dysfunction and 17 OTHER. An example vignette for each category of classification can be found in Appendix E - H.

3.2.3 Phase 2

3.2.3.1 Participants

For phase two, six different raters were recruited based on previous willingness to participate in research from the publicly available list of MDT practitioners registered with McKenzie Institute International who practice in Canada or the United States. They were required to classify the 53 clinical vignettes generated in phase one of the study. The sample size of six raters was chosen to provide equal division of groups by level of MDT training and within group variability. To be included, the rater had to be a Credentialed or Diploma with the McKenzie Institute and have applied the MDT system to the extremities for more than three years. Raters were excluded if they participated in the creation of the clinical vignettes, did not wish to participate, were unable to understand written and spoken English, unable to provide informed consent or were unable to follow the instructions for rating the clinical vignettes. Correspondence was conducted and informed consent was obtained from each rater through electronic mail (Appendix I to J). In total, all six raters recruited consented to participation in this study.

3.2.3.2 Data Collection and Procedures

Demographic information like gender and age were collected along with other relevant characteristics like clinical practice setting, years of practice, length of time Credentialed/Diploma, discipline (eg. Physiotherapist versus Doctor of Chiropractic), proportion of extremity patients treated and proportion of knee patients treated with MDT. Data collection forms can be found in Appendix J. For each vignette, the raters were instructed to review the vignette and based on the history and clinical presentation, assign the vignette a classification of Derangement, Articular Dysfunction, Contractile Dysfunction or OTHER. Each vignette was randomly assigned a number from 1 to 53 to facilitate tracking of responses and data collection. All raters were blinded to the provisional MDT classification originally assigned to the vignette by its creator in phase one.

3.3 Analysis

Descriptive statistics for the demographic and clinical information for the raters were determined. Inter-rater reliability, our primary objective, was determined through the calculation of Fleiss kappa statistic along with 95% confidence interval (CI) and standard error (SE) across all six raters for all categories (Fleiss, 1971; Fleiss, Nee & Landis, 1979). Data were analyzed for Fleiss kappa using a program written in Matlab version 7.14 (Cardillo, 2007). Kappa values were interpreted using definitions outlined by Landis and Koch (1977): 0.01 to 0.20 slight agreement, 0.21 to 0.40 fair agreement, 0.41 to 0.60 moderate agreement, 0.61 to 0.80 substantial agreement and 0.81 to 1.00 almost perfect agreement.

It has been suggested that although overall kappa for three or more raters may lead to a better representation of reliability, overall kappa may mask extreme cases of agreement or disagreement for paired raters (O'Leary et al., 2014). A solution is to report both overall and paired kappa data to provide the most informative summary. Paired comparisons of the agreement of vignette classification among the six raters were thus analyzed and reported as percentage agreement and kappa statistic with standard error. Also, frequency distribution of the category of classification was analyzed for each

individual rater and reported as a whole number and percentage of total number of vignettes. Additionally, individual raters and their agreement with the vignette provisional classification assigned in phase one were analyzed and reported as percentage agreement and kappa statistic with standard error. Raters were grouped based on their level of education.

To examine if the level of education influenced the reliability, differences in Fleiss kappa values between Credentialed and Diploma therapists were compared. A bootstrap method with a 1000 samples was utilized and Fleiss kappa coefficients were calculated separately for the Credential and Diploma raters for each of these samples (McKenzie et al., 1996). The differences between the Fleiss kappa coefficients were determined. The mean of these differences was determined along with the 95% confidence interval represented by the 25 and 975 values. If the 95% confidence interval included zero, then no significant difference existed between the Credential and Diploma raters.

3.4 Results

The six raters recruited to rate the clinical vignettes were all physiotherapists and comprised of three Credentialed and three Diploma therapists. Four raters practiced fee for service and two worked in multiple settings. Four raters were male and two were female. Demographic information obtained from each of the raters is displayed in Table 3.

The overall kappa value amongst the six raters demonstrated substantial agreement with kappa=0.72 (SE=0.02) with a 95% CI of 0.71 to 0.73. The highest level of reliability was for the Derangement category with kappa=0.83; the lowest level of reliability was for the OTHER category with kappa=0.64. Articular and Contractile Dysfunction had a kappa of 0.67 and 0.69, respectively. There was 100% agreement in classification among all six raters in 31 of the 53 (58.5%) clinical vignettes.

Table 3 Demographic information of participating phase 2 raters (n=6)

Variables	Mean (SD)	Range
Age, years	51 (13.4)	35 to 67
Years of Practice	25 (13.7)	10 to 44
Proportion of Extremity Patients Treated in Practice, percentage	37 (16.0)	20 to 65%
Proportion of Knee Patients of Peripheral Joints in Practice, percentage	79 (37.2)	5 to 100%

SD - standard deviation

The frequency distribution of the category of classification by individual raters is displayed in Table 4. Derangement syndrome was the mostly commonly assigned classification to the vignettes across all raters, ranging from 20 (38%) – 26 (49%) of the total number (53) of vignettes reviewed.

Paired comparison of agreement in vignette classifications across the six raters are displayed in Table 5. The top right half shows percentage agreement and the bottom left half shows kappa scores (standard error) for all possible pairings of raters. The highest percentage of agreement (92%) and kappa (0.89) were between rater 1 and 5. The lowest percentage of agreement (72%) and kappa (0.58) were between rater 2 and 6. Reliability between raters showed moderate to substantial agreement.

Individual rater responses were compared to the provisional classifications of the vignettes and reliability calculated with results displayed in Table 6. Raters are grouped by level of MDT training. The highest percentage of agreement with the provisional classification was 91% for rater 5 and the lowest agreement was 81% for rater 6. Kappa scores ranged from 0.73 for rater 6 and 0.86 for rater 5. All raters showed substantial reliability (kappa ≥ 0.73) with the provisional classification.

Reliability by MDT training across raters for Credentialed and Diploma therapists are shown in Tables 7. The mean difference value between kappa values for Credentialed and Diploma therapists was -0.03 (95% CI -0.15 to 0.11). Since the confidence interval includes 0, there is no significant difference between rater groups based on level of education.

Table 4 Frequency distribution of category of classification by individual rater

Rater	Classification n (%)				Total
	Derangement	Articular Dysfunction	Contractile Dysfunction	OTHER	
1	25 (47%)	6 (11%)	5 (9%)	17 (32%)	53
2	22 (41%)	6 (11%)	6 (11%)	19 (36%)	53
3	24 (45%)	7 (13%)	8 (15%)	14 (26%)	53
4	20 (38%)	11 (21%)	8 (15%)	14 (26%)	53
5	25 (47%)	7 (13%)	6 (11%)	15 (28%)	53
6	26 (49%)	5 (9%)	12 (23%)	10 (19%)	53

n - number of vignettes

Table 5 Percentage agreement and kappa (standard error) for paired comparisons among the six raters

Percentage Agreement						
Rater	1	2	3	4	5	6
1	-	83%	89%	77%	92%	77%
2	0.74 (0.09)	-	75%	75%	81%	72%
3	0.83 (0.09)	0.64 (0.09)	-	81%	91%	79%
4	0.67 (0.08)	0.65 (0.08)	0.73 (0.08)	-	83%	75%
5	0.89 (0.09)	0.72 (0.09)	0.86 (0.09)	0.76 (0.08)	-	81%
6	0.66 (0.09)	0.58 (0.09)	0.72 (0.08)	0.65 (0.08)	0.72 (0.09)	-
Kappa (Standard Error)						

Table 6 Percentage agreement and kappa (standard error) of individual raters versus the provisional classification grouped by MDT education

Rater	Statistic	Diploma			Credentialed		
		1	2	3	4	5	6
Provisional Classification	% Agreement	88%	88%	85%	83%	91%	81%
	Kappa	0.83	0.83	0.78	0.76	0.86	0.73
	(SE)	(0.09)	(0.09)	(0.09)	(0.08)	(0.09)	(0.08)

SE – standard error

Table 7 Reliability by MDT education across raters

MDT Education	Kappa (SE)	95% CI
Credentialed (n=3)	0.71 (0.048)	0.61 to 0.80
Diploma (n=3)	0.74 (0.051)	0.64 to 0.84

SE – standard error; CI – confidence interval

Chapter 4

4 Discussion

This chapter reviews the key findings of the research study and discusses the implications of these results for clinicians. Limitations of the study and recommendations for future research are also outlined.

4.1 Overview

The primary objective of this thesis was to determine the inter-rater reliability of the McKenzie System of MDT when trained therapists classify musculoskeletal knee pain using patient based clinical vignettes. The lack of research on the clinical utility of the MDT system when it is applied in the extremities, and specifically the knee, was the motivation for this thesis project.

4.2 Key Findings of the Thesis Project

The primary findings of this study suggest that the inter-rater reliability of Credentialed and Diploma clinicians within the MDT Institute demonstrate "substantial agreement" when using the MDT system to classify patients presenting with musculoskeletal knee pain (kappa=0.72). There was no statistically significant difference between Credentialed or Diploma raters (CI -0.15 to 0.11). Thus, it appears that clinicians with specific MDT training can use the MDT system to assess and classify patients with knee pain using clinical vignettes.

The results of this study are consistent with others evaluating the use of MDT in the extremity (Heidar Abady et al., 2014; Kelly, May, & Ross, 2008; May & Ross, 2009) and spine (Clare et al. 2005; Kilpikoski et al. 2002; Razmjou et al., 2000). The reliability of paired raters for the current study was kappa=0.58 to 0.89. Similarly, Razmajou et al. (2000) found overall reliability of kappa=0.70 when investigated the inter-rater reliability of two MDT trained examiners, one Diploma and one Credentialed, assessing real patients presenting with mechanical low back pain. Likewise, Kilpikoski et al. (2002) found an overall reliability between two MDT trained Diploma raters performing

independent, consecutive assessments of 39 patients with low back pain to be moderate (kappa=0.6). Using 14 raters, 7 Credentialed and 7 Diploma, Clare et al. (2005) reported the overall reliability for the classification of patients with lumbar and cervical pain was almost perfect (kappa=0.84) among paired raters. Hence, the MDT assessment appears to be a reliable method of assessment for patients presenting with musculoskeletal spinal or knee pain.

In the current study, individual raters demonstrated substantial reliability (kappa=0.73 to 0.86) while percentage of agreement ranged from 81 to 91% when rater classification was compared to the vignette provisional classification. Methodologically similar to the current study, Heidar Abady et al. (2014) used 54 clinical vignettes of patients presenting with musculoskeletal shoulder disorders to evaluate the inter-reliability of the MDT assessment by six Diploma clinicians and reported reliability of kappa=0.89 (0.77 to 0.96) and 95% overall agreement across raters against the provisional classification. Thus, MDT seems to have similar reliability when classifying clinical vignettes of patients presenting with musculoskeletal knee and shoulder pain.

The highest level of reliability was for the Derangement category with kappa=0.83 while the level of reliability for the three remaining categories varied with kappa=0.62 to 0.69. The difference between these levels of reliability may be explained to some degree by the general presentation of each category. The Derangement syndrome, by definition, is readily identifiable by a lasting reduction or elimination of patients' symptoms through repeated movement in a particular direction (McKenzie and May, 2003). The relative lower level of agreement of the remaining categories may be partially attributable to the absence of a unique identifiable characteristic, such as directional preference that is present with Derangement syndrome. There may also be less familiarity with the extremity classifications and their criteria which have been more recently defined in the literature (May and Rosedale, 2012). This is particularly true for the OTHER category as multiple subgroups are included which makes determining a classification based exclusively on an initial assessment more challenging.

The relative reliability and similarity of kappa values between Credentialed (kappa=0.71) and Diploma (kappa=0.74) clinicians was anticipated. When grouped by education, Diploma raters compared to the provisional classification demonstrated reliability of kappa=0.78 to 0.83 and Credentialed raters reliability of kappa=0.73 to 0.86. Although Diploma holders of the MDT Institute undergo further education, each clinician in our study had a great deal of experience using the extremity assessment form and in treating patients with musculoskeletal pain in the extremities, and specifically, the knee. Thus, varying degrees of MDT competency did not appear to negatively impact the overall level of reliability of the system when evaluating musculoskeletal knee pain. However, we did not include raters without Credentialed or Diploma competence and as a result the effect of lower levels of MDT training was not evaluated.

4.3 Clinical Implications

It has been suggested that an assessment to simply identify structures as the cause of pain does not elicit enough information to understand the problem or to justify a course of management (Jones and Rivett, 2004). Indeed, specific features of the current model of examination for musculoskeletal knee pain, medical imaging and OSTs, have cast doubt on the clinical utility of this model. OSTs have demonstrated questionable diagnostic accuracy to discriminate the anatomical structures they are said to identify and medical imaging has brought to light the confounding prevalence of pathology and abnormalities in asymptomatic individuals. These findings would suggest a model less reliant on an anatomical diagnosis may be worth evaluating.

It has also been suggested that classification systems like MDT, may offer better clinical utility as the categories of classification are based on patient's responses to repeated mechanical loading strategies rather than the presence of patho-anatomy (Lynch & May, 2013; May & Rosedale, 2012; Rosedale et al., 2014). An MDT assessment directs treatment with an appropriate loading strategy and in the presence of a directional preference, may determine who might and might not respond to treatment (Rosedale et al., 2014). An MDT assessment has shown to have good reliability in the spine and extremity, and while promising, more work needs to be done around the efficacy of MDT guided interventions.

There are two studies within the literature addressing inter-rater reliability of MDT in musculoskeletal pain in the lower extremity. (Kelly, May & Ross, 2008; May & Ross, 2009) To our knowledge, this is the first study to investigate the reliability of the MDT specifically for musculoskeletal knee pain alone. The results of this study on the knee reinforce the findings of previous reliability studies, indicating that the McKenzie System of MDT appears to be a reliable approach to assessing musculoskeletal pain in the knee.

Two previous studies have examined the use of an MDT guided intervention in the treatment of knee pain. Lynch & May (2013) documented the case study of directional preference of the knee using MDT. Although presenting with a positive McMurray's test and pain with swimming, following prescribed exercise matching the directional preference, the patient reported 95% improvement in function and symptoms and a negative McMurray's test. It was concluded by the authors that the use of McMurray's test was not diagnostic and the result of the test appeared irrelevant and only clinical useful as a symptomatic baseline. More recently, Rosedale et al., (2014) published a randomized control trial using an MDT guided assessment to determine a directional preference for patients with end stage OA. Patients were classified as either Derangement or no Derangement. Patients who were matched with exercises consistent with a directional preference demonstrated significant decreases in pain and increases in self-report function scales after 2 weeks with large effect sizes (d=0.98 to 1.44). Although the effect sizes decreased at 3 months, they remained small to large (d=0.42 to 0.80) compared to the control group and patients without a directional preference. While the results from the study cannot be directly attributed to the MDT classification, it was concluded that the response to directional preference matched exercises should be explored further. With this in mind, it is reasonable to speculate that the method of MDT classification could facilitate the identification of who will or will not respond to treatment which in turn may enable the clinician to match the most appropriate treatment to various patient subgroups.

4.4 Limitations

Raters used in this study have achieved a high level of education and understanding of the McKenzie System of MDT and have significant experience applying the system to

musculoskeletal problems in the extremity. The background of the raters and subsequent findings will limit the generalizability of the results to those individuals with similar training and experience. As a result, generalizing the findings to practitioners without this level of training may not be appropriate. The vignettes with the provisional classification of OTHER were not further disseminated into subgroups for the raters to identify. To do so would have increased the number of categories of classification from 4 to 13, thus requiring an increase in the number of clinical vignettes for raters to review. As such, the reliability of raters classifying patients into those subgroups and the direction of subsequent treatment cannot be determined from this study. Additionally, the reviewers of each of the vignettes (SW, RR) were not blinded to the creator of the vignette or the provisional classification assigned to each vignette. This could result in the creation of a biased sample. Another potential limitation is the use of clinical vignettes as an alternative to real patients. It has been argued that vignettes cannot measure correspondence of hypothetical behavior and real world behavior (Evans et al., 2015) and may not capture the subtlety of a patients' presentation, oversimplify findings making a diagnosis easier and potentially inflating calculated agreement (Peabody et al., 2000; Werneke, Hart, Deutscher and Stratford, 2011). However, clinical vignettes offer the convenience of collecting information from multiple sources simultaneously while allowing for the flexibility of variable manipulation to ensure a heterogeneous sample.

4.5 Future Recommendations

This study found "substantial agreement" among Credentialed and Diploma holders in MDT. To generalize the use of the system to more users, future research should continue to investigate the reliability of MDT using raters with lower levels of training and experience.

Although the results of the study are encouraging, the results are limited to the reliability of raters classifying knee pain using clinical vignettes. To generalize the system further, future research should be conducted with real patients to demonstrate reliability in a real world, clinical setting.

Further to this, the efficacy of MDT guided treatment for patients presenting with musculoskeletal knee pain should be explored further. Long, Donelson & Fung (2004) found that exercises matching subjects' directional preference in the lumbar spine significantly decreased pain and improved primary outcomes. The effect of directional preference matched exercises as indicated by the MDT classification needs to be evaluated and measured to substantiate use for clinical intervention in musculoskeletal knee pain.

4.6 Conclusion

In conclusion, this is the first study investigating the inter-rater reliability of the McKenzie System of MDT in the examination of musculoskeletal knee pain. The McKenzie System of MDT demonstrated substantial agreement, indicating acceptable inter-rater reliability for trained raters when using clinical vignettes to classify patients presenting with musculoskeletal knee pain. The results of this study offer preliminary support for the use of MDT in the assessment of musculoskeletal knee pain and support for future studies.

References

Abady, A. H., Rosedale, R., Overend, T. J., Chesworth, B. M., & Rotondi, M. A. (2014). Inter-examiner reliability of diplomats in the mechanical diagnosis and therapy system in assessing patients with shoulder pain. Journal of Manual & Manipulative Therapy, 22(4), 199–205.

Aina, A., & May, S. (2005). A shoulder derangement. Manual Therapy, 10(2), 159–63.

Atzmüller, C., & Steiner, P. M. (2010). Experimental Vignette Studies in Survey Research. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 6(3), 128–138.

Beattie, K. B, Boulos, P., Pui, M., O'Neill, J., Inglis, D., Webber, C. E., & Adachi, J. D. (2005). Abnormalities identified in the knees of asymptomatic volunteers using peripheral magnetic resonance imaging. Osteoarthritis and Cartilage, 13(3), 181–6.

Boks, S. S., Vroegindeweij, D., Koes, B. W., Hunink, M. M. G. M., & Bierma-Zeinstra, S. M. A. (2006). Magnetic resonance imaging abnormalities in symptomatic and contralateral knees: prevalence and associations with traumatic history in general practice. The American Journal of Sports Medicine, 34(12), 1984–91.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, Prevalence and Kappa. Journal of Clinical Epidemiology, 46(5), 423–429.

Campbell, J., Harte, A., Kerr, D. P., & Murray, P. (2014). The location of knee pain and pathology in patients with a presumed meniscus tear: preoperative symptoms compared to arthroscopic findings. Irish Journal of Medical Science, 183(1), 23–31.

Cardillo G. (2007) Fleiss'es kappa: compute the Fleiss'es kappa for multiple raters. http://www.mathworks.com/matlabcentral/fileexchange/15426

Cicchetti, D. V. & Feinstein, A. R. (1990). High agreement but low kappa: II. Resolving the paradoxes. Journal of Clinical Epidemiology, 43(6), 551–558.

Clare, H. A, Adams, R., & Maher, C. G. (2005). Reliability of McKenzie classification of patients with cervical or lumbar pain. Journal of Manipulative and Physiological Therapeutics, 28(2), 122–7.

Cook, C. (2010). The lost art of the clinical examination: an overemphasis on clinical special tests. The Journal of Manual & Manipulative Therapy, 18(1), 3–4.

Cook, C., Mabry, L., Reiman, M. P., & Hegedus, E. J. (2012). Best tests/clinical findings for screening and diagnosis of patellofemoral pain syndrome: a systematic review. Physiotherapy, 98(2), 93–100.

de Vet, H. C. W., Mokkink, L. B., Terwee, C. B., Hoekstra, O. S., & Knol, D. L. (2013). Clinicians are right not to like Cohen's κ. BMJ (Clinical Research Ed.), 346, 1-7.

Dresselhaus, T. R., Peabody, J. W., Luck, J., & Bertenthal, D. (2004). An evaluation of vignettes for predicting variation in the quality of preventive care. Journal of General Internal Medicine, 19(10), 1013–8.

Evans, S. C., Roberts, M. C., Keeley, J. W., Blossom, J. B., Amaro, C. M., Garcia, A. M., Stough C.O., Canter, K.S., Robles, R., & Reed, G. M. (2015). Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. International Journal of Clinical and Health Psychology, 15(2), 160–170.

Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I. The Problems of Two Paradoxes. Journal of Clinical Epidemiology, 43(6), 543–549.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. Psychological Bulletin, 76(5), 378–382.

Fleiss, J. L., Nee, J. C., & Landis, J. R. (1979). Large sample variance of kappa in the case of different sets of raters. Psychological Bulletin, 86(5), 974–977.

Gage, B. E., McIlvain, N. M., Collins, C. L., Fields, S. K., & Comstock, R. D. (2012). Epidemiology of 6.6 million knee injuries presenting to United States emergency departments from 1999 through 2008. Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine, 19(4), 378–85.

Geraets, S. E. W., Meuffels, D. E., van Meer, B. L., Breedveldt Boer, H. P., Bierma-Zeinstra, S. M. A, & Reijman, M. (2015). Diagnostic value of medical history and physical examination of anterior cruciate ligament injury: comparison between primary care physician and orthopaedic surgeon. Knee Surgery, Sports Traumatology, Arthroscopy, 23, 968–974.

Goossens, P., Keijsers, E., van Geenen, R. J. C., Zijta, A., van den Broek, M., Verhagen, A. P., & Scholten-Peeters, G. G. M. (2015). Validity of the Thessaly Test in Evaluating Meniscal Tears Compared With Arthroscopy: A Diagnostic Accuracy Study. Journal of Orthopaedic & Sports Physical Therapy, 45(1), 18–24.

Gould, D. (1996). Using Vignettes to collect data for nursing research studies: how valid are the findings? Journal of Clinical Nursing, 5, 207–212.

Gupta, S., Hawker, G. A., Laporte, A., Croxford, R., & Coyte, P. C. (2005). The economic burden of disabling hip and knee osteoarthritis (OA) from the perspective of individuals living with this condition. Rheumatology, 44(12), 1531–1537.

Hegedus, E. J., Cook, C., Hasselblad, V., Goode, A., & McCrory, D. C. (2007). Physical examination tests for assessing a torn meniscus in the knee: a systematic review with meta-analysis. The Journal of Orthopaedic and Sports Physical Therapy, 37(9), 541–50.

Jinks, C., Jordan, K., & Croft, P. (2007). Osteoarthritis as a public health problem: The impact of developing knee pain on physical function in adults living in the community: (KNEST 3). Rheumatology, 46(5), 877–881.

Jones, M. A., & Rivett, D. A. (2004). Introduction to clinical reasoning. In Clinical Reasoning for Manual Therapists (pp. 16–17). Edinburgh: Butterworth Heinemann.

Kaplan, L. D., Schurhoff, M. R., Selesnick, H., Thorpe, M., & Uribe, J. W. (2005). Magnetic resonance imaging of the knee in asymptomatic professional basketball players. Arthroscopy: The Journal of Arthroscopic & Related Surgery : Official Publication of the Arthroscopy Association of North America and the International Arthroscopy Association, 21(5), 557–61.

Kelly, E., May, S., & Ross, J. (2008). Original Research The reliability of the McKenzie classification system using extremity McKenzie assessment forms. International Journal of Mechanical Diagnosis and Therapy, 3(3), 3–6.

Kilpikoski, S., Airaksinen, O., Kankaanpää, M., Leminen, P., Videman, T., & Alen, M. (2002). Interexaminer reliability of low back pain assessment using the McKenzie method. Spine, 27(8), 207–14.

Kirkley, A., Birmingham, T. B., Litchfield, R. B., Giffin, J. R., Willits, K. R., Wong, C. J., Feagan, B.G., Donner, A., Griffin S.H., D'Ascanio, L.M., Pope, J.E., & Fowler, P. J. (2008). A randomized trial of arthroscopic surgery for osteoarthritis of the knee. The New England Journal of Medicine, 359(11), 1097–1107.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1), 159–174.

Lange, T., Freiberg, A., Dröge, P., Lützner, J., Schmitt, J., & Kopkow, C. (2015). The reliability of physical examination tests for the diagnosis of anterior cruciate ligament rupture – A systematic review. Manual Therapy, 20, 402-411

Laprade, R. F., Quinter, B. M., Veenstra, M. A., & Hodgman, C. G. (1994). The Prevalence of Abnormal Magnetic Resonance Imaging Findings in Asymptomatic Knees. The American Journal of Sports Medicine, 22(6), 739–745.

Leblanc, M.-C., Kowalczuk, M., Andruszkiewicz, N., Simunovic, N., Farrokhyar, F., Turnbull, T. L., Debski, R.E., & Ayeni, O. R. (2015). Diagnostic accuracy of physical examination for anterior knee instability: a systematic review. Knee Surgery, Sports Traumatology, Arthroscopy. March 13.

Lijmer, J. G., Mol, B.W., Heisterkemp, S., Bonsel, G.J., Prins, M.H., van der Meulen, J.H.P., & Bossuyt, P.M.M. (1999). Empirical Evidence of Design-Related Bias, 282(11).

Littlewood, C., & May, S. (2007). A contractile dysfunction of the shoulder. Manual Therapy, 12(1), 80–3.

Long, A., Donelson, R., & Fung, T. (2004). Does it matter which exercise? A randomized control trial of exercise for low back pain. Spine, 29(23), 2593–602.

Lynch, G., & May, S. (2013). Directional preference at the knee: a case report using mechanical diagnosis and therapy. The Journal of Manual & Manipulative Therapy, 21(1), 60–6.

MatLab, Version 7.14. (2012). MathWorks, Natick, MA, USA

May, S., & Aina, A. (2012). Centralization and directional preference: a systematic review. Manual Therapy, 17(6), 497–506.

May, S. J., & Rosedale, R. (2012). A survey of the McKenzie Classification System in the Extremities: prevalence of mechanical syndromes and preferred loading strategies. Physical Therapy, 92(9), 1175–86.

May, S., & Ross, J. (2009). The McKenzie classification system in the extremities: a reliability study using McKenzie assessment forms and experienced clinicians. Journal of Manipulative and Physiological Therapeutics, 32(7), 556–63.

McKenzie, D. P., Mackinnon, A. J., Péladeau, N., Onghena, P., Bruce, P. C., Clarke, D., Harrigan, S., & M., McGorry, P. D. (1996). Comparing correlated kappas by resampling: Is one level of agreement significantly different from another? Journal of Psychiatric Research, 30(6), 483–492.

McKenzie, R., & May, S. (2000). The Human Extremities: Mechanical Diagnosis and Therapy. Spinal Publications New Zealand Ltd.

McKenzie, R., & May, S. (2003). The Lumbar Spine: Mechanical Diagnosis and Therapy (Second Edition). Spinal Publications New Zealand Ltd.

Nguyen, U.D.T., Zhang, Y., Yanyan, Z., Niu, J., Zhang, B., Aliabadi, P., & Felson, D. (2011). Increasing Prevalence of Knee Pain and Symptomatic Knee Osteoarthritis. Ann Intern Med, 155(11), 725–732.

O'Leary, S., Lund, M., Ytre-Hauge, T. J., Holm, S. R., Naess, K., Dalland, L. N., & McPhail, S. M. (2014). Pitfalls in the use of kappa when interpreting agreement between multiple raters in reliability studies. Physiotherapy (United Kingdom), 100(1), 27–35.

Peabody, J. W., Luck, J., Glassman, P., Dresselhaus, T. R., & Lee, M. (2000). Comparison of Vignettes, Standardized Patients, and Chart Abstraction. The Journal of American Medical Association, 283(13), 1715–1722.

Peabody, J. W., Luck, J., Glassman, P., Jain, S., Hansen, J., Spell, M., & Lee, M. (2004). Measuring the Quality of Physician Practice by Using Clinical Vignettes: A Prospective Validation Study. Annals of Internal Medicine, 141, 771–780.

Peeler, J., Leiter, J., & MacDonald, P. (2010). Accuracy and reliability of anterior cruciate ligament clinical examination in a multidisciplinary sports medicine setting. Clinical Journal of Sport Medicine, 20(2), 80–5.

Razmjou, H., Kramer, E., & Yamada, R. (2000). Intertester Reliability of the McKenzie Evaluation in Assessing Patients With Mechanical Low-Back Pain. Journal of Orthopaedic & Sports Physical Therapy, 30(7), 368–389.

Rosedale, R., Rastogi, R., May, S., Chesworth, B. M., Filice, F., Willis, S., Howard, J., Naudie, D., & Robbins, S. M. (2014). Efficacy of Exercise Intervention as Determined by the McKenzie System of Mechanical Diagnosis and Therapy for Knee Osteoarthritis: A Randomized Controlled Trial. The Journal of Orthopaedic and Sports Physical Therapy, 44(3).

R Project for Statistical Computing, Version 3.1.3 (2014). Vienna, Austria.

Rotondi, M. A., & Donner, A. (2012). A confidence interval approach to sample size estimation for studies of interobserver agreement with multiple raters and outcomes. Journal of Clinical Epidemiology, 65(7), 778-84.

Rotondi, M.A. (2013). Sample Size Estimation Functions for Studies of Interobserver Agreement. http://cran.r-project.org/web/packages/kappaSize/index.html

Rutten, G. M. J., Harting, J., Rutten, S. T. J., Bekkering, G. E., & Kremers, S. P. J. (2006). Measuring physiotherapists' guideline adherence by means of clinical vignettes: a validation study. Journal of Evaluation in Clinical Practice, 12(5), 491–500.

Saleh, M. N., Korner-Bitensky, N., Snider, L., Malouin, F., Mazer, B., Kennedy, E., & Roy, M. A. (2008). Actual vs. best practices for young children with cerebral palsy: a survey of paediatric occupational therapists and physical therapists in Quebec, Canada. Developmental Neurorehabilitation, 11(1), 60–80.

Sihvonen, R., Paavola, M., Malmivaara, A., Itälä, A., Joukainen, A., Nurmi, H., Kalske, J., & Järvinen, T. L. N. (2013). Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. The New England Journal of Medicine, 369(26), 2515–24.

Thorlund, J. B., Juhl, C. B., Roos, E. M., & Lohmander, L. S. (2015). Arthroscopic surgery for degenerative knee: systematic review and meta-analysis of benefits and harms. BMJ, 350, 1–9.

Veloski, J., Tai, S., Evans, A. S., & Nash, D. B. (2005). Clinical vignette-based surveys: a tool for assessing physician practice variation. American Journal of Medical Quality, 20(3), 151–7.

Werneke, M. W., Deutscher, D., Hart, D. L., Stratford, P., Ladin, J., Weinberg, J., Herbowy, S., & Resnik, L. (2014). McKenzie lumbar classification: inter-rater agreement by physical therapists with different levels of formal McKenzie postgraduate training. Spine, 39(3), E182–90.

Werneke, M. W., Hart, D. L., Deutscher, D., & Stratford, P. W. (2011). Clinician's ability to identify neck and low back interventions: an inter-rater chance-corrected agreement pilot study. Journal of Manual & Manipulative Therapy, 19(3), 172–181.

Zanetti, M., Pfirrmann, C. W. A, Schmid, M. R., Romero, J., Seifert, B., & Hodler, J. (2003). Patients with Suspected Meniscal Tears: Prevalence of Abnormalities Seen. American Journal of Radiology, 181, 635–641.

Appendices

Appendix A Ethics Approval from Western University Health Science Research Ethics Board

Western Research

Research Ethics

Western University Health Science Research Ethics Board
HSREB Delegated Initial Approval Notice

Principal Investigator: Dr. Trevor Birmingham
Department & Institution: Health Sciences\Physical Therapy,Western University

HSREB File Number: 105498
Study Title: Inter-observer Reliability of the McKenzie System of Mechanical Diagnosis and Therapy in the Examination of the Knee
Sponsor:

HSREB Initial Approval Date: August 26, 2014
HSREB Expiry Date: July 31, 2015

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Revised Western University Protocol	Amendment to Western Protocol	2014/08/08
Amendment	Amendment to Data Collection Form	2014/08/08
Amendment	Amendment to Lower Extremity Assessment Form	2014/08/04
Revised Letter of Information & Consent	Amendments to Phase 1 LOI	2014/08/08
Revised Letter of Information & Consent	Amendments to LOI Phase 2	2014/08/08

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above named study, as of the HSREB Initial Approval Datenoted above.

HSREB approval for this study remains valid until the HSREB Expiry Date noted above, conditional to timely submission and acceptance of HSREB ContinuingEthics Review. If an Updated Approval Notice is required prior to the HSREB Expiry Date, the Principal Investigator is responsible for completing andsubmitting an HSREB Updated Approval Form in a timely fashion.

The Western University HSREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), theInternational Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Guideline for Good Clinical PracticePractices (ICH E6 R1), the Ontario Personal Health Information Protection Act (PHIPA, 2004), Part 4 of the Natural Health Product Regulations, Health CanadaMedical Device Regulations and Part C, Division 5, of the Food and Drug Regulations of Health Canada.

Members of the HSREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when theyare presented to the REB.

The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Appendix B Phase 1 Recruitment Email

Dear MDT Clinician,

You are invited to participate in a research study that is investigating the inter-examiner reliability of the McKenzie System of MDT when used by MDT trained clinicians in patients presenting with musculoskeletal knee pain. For this study, 10 clinicians are required to create 53 clinical vignettes for 6 expert MDT raters to review and classify.

We are asking you because you are Credentialed and/or have a Diploma standing within the McKenzie Institute and have been applying MDT to musculoskeletal problems in the extremity for greater than 3 years.

This study will be conducted by Trevor Birmingham, a Professor in the School of Physical Therapy at Western University in the School of Physical Therapy. Sean Willis, a Master of Science student in Health and Rehabilitation Sciences at Western University, Shawn Robbins, an Assistant Professor in the School of Physical and Occupational Therapy at McGill University and Richard Rosedale, an Instructor with the McKenzie Institute will also be participating in the study.

Attached is a Letter of Information for you to review and consider your participation. Also attached is a Consent Form for you to sign and submit should you wish to participate in this study.

Thank you for taking the time to consider participation in this study.

Sincerely,

Sean Willis

Appendix C Phase 1 Letter of Information and Consent Form

School of Physical Therapy

Letter of Information and Consent Form

Clinical Vignette Creation

Inter-observer Reliability of the McKenzie System of Mechanical Diagnosis and Therapy in the Examination of the Knee

Study Principal Investigator
Trevor Birmingham PhD, BSc(PT)
Western University, School of Physical Therapy

Study Co-Investigators
Sean Willis BSc(PT), MSc Student
Western University, Health and Rehabilitation Sciences

Shawn Robbins PhD, BSc(PT)
McGill University, School of Physical and Occupational Therapy

Study Advisor
Richard Rosedale PT, Dip MDT
Instructor, McKenzie Institute

1. Introduction

You are invited to participate in a research study that is investigating the inter-examiner reliability of the McKenzie System of Mechanical Diagnosis and Therapy (MDT) when used by MDT trained clinicians in patients presenting with musculoskeletal knee pain. For this study, 10 clinicians are required to create 53 clinical vignettes for 6 expert MDT raters to review and

classify. The purpose of this letter is to provide you with the necessary background and information to make an informed decision regarding your participation within this study.

2. Why Are We Asking You?

We are asking you because you are Credentialed and/or have a Diploma standing within the McKenzie Institute and have been applying MDT to musculoskeletal problems in the extremity for greater than 3 years.

We are giving this letter of information only to people who are registered on the publically available list of MDT clinicians practicing in Canada and the United States and who understand both written and spoken English. If the above situation does not apply to you, we ask that you do not volunteer to participate in our study.

3. Who is Conducting the Study?

This study will be conducted by Trevor Birmingham, a Professor in the School of Physical Therapy at Western University in the School of Physical Therapy. Sean Willis, a Master of Science student in Health and Rehabilitation Sciences at Western University, Shawn Robbins, an Assistant Professor in the School of Physical and Occupational Therapy at McGill University and Richard Rosedale, an Instructor with the McKenzie Institute will also be participating in the study.

4. What will happen if you agree?

If you agree to participate, Sean Willis will contact you via email with a blank McKenzie Extremity Assessment form and instructions on generating clinical vignettes based on past patient assessment files. You will be asked to review your files and transfer the appropriate clinical information on the forms provided. This information should include a subjective history of the presenting musculoskeletal issue, signs and symptoms and relevant past medical history. The vignette should also include objective findings pertaining to baseline ROM, strength and neurological findings, functional baseline activities and responses to selected repeated movements. Each of the vignettes submitted are to be de-identified to only include age range (eg. 35-40 years old), gender and a category of occupation to respect the confidentiality of your patients. You are asked to provide clinical vignettes that are characteristic of the following four mechanical syndromes: (2) derangement, (2) other, (1) contractile dysfunction and (1) articular dysfunction.

It is expected that the review of your past patient assessments and creation of the vignettes will take some time. The anticipated time commitment for the creation of the clinical vignettes is 1 hour per vignette for a total of 6 hours. 3 months will be provided to complete this process.

5. What are the risks and benefits?

There are no direct benefits from participating in this study. However, for clinicians who treat people with musculoskeletal knee disorders, this would be the first study to address inter-examiner reliability of McKenzie system in assessing patients with knee pain. It would provide MDT practitioners with valuable information about reliability of this assessment method.

There are no risks to participating in this study.

6. What about confidentiality?

Your confidentiality will be protected. No information that discloses your identity will be released or be a part of any publication of the results of the study. Your first and last name as well as your email address will be collected for contact and communication purposes only. All records will be given a code ID to be used on all collected forms. All information collected will be kept in locked filing cabinets and shredded seven years after the completion of the study.

7. Alternatives to participation?

Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions or withdraw from the study at any time with no effect on your future employment. Should you withdraw from the study, you may also withdraw any vignettes that you have submitted.

Initials _____

8. If you agree to participate in the study

Please sign the attached consent form and return to Sean Willis. You do not waive any of your legal rights by signing the consent form.

You may keep this letter of information. A copy of your signed consent form will be made for you. If you have any questions about this study, please contact Sean Willis.

If you have any questions about your rights as a research participant or the conduct of the study, you may contact The Office of Research Ethics at Western University.

Representatives of The University of Western Ontario Health Sciences Research Ethics Board may contact you or require access to your study-related records to monitor the conduct of the research.

Western HealthSciences

School of Physical Therapy

Consent Form

Inter-observer Reliability of the McKenzie System of Mechanical Diagnosis and Therapy in the Examination of the Knee

I have read the accompanying letter of information and I agree to participate. Questions I had have been answered to my satisfaction.

Date	Participant's Name	Participant's Signature
_____	_____	_____

Date	Name of Person Obtaining Informed Consent	Signature of Person Obtaining Informed Consent
_____	_____	_____

Initials _____

Appendix D McKenzie Institute Lower Extremities Assessment Form

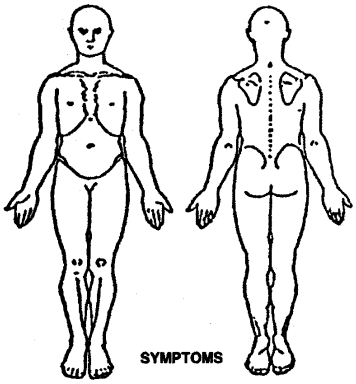
THE McKENZIE INSTITUTE
LOWER EXTREMITIES ASSESSMENT

Vignette ID # : _____

Age Range (eg. 35-45 yrs.) : _____

Gender : M F

Category of Occupation : _____



Work: Mechanical stresses _____

Leisure: Mechanical stresses _____

Functional disability from present episode _____

Functional disability score _____

VAS Score (0-10) _____

HISTORY

Present symptoms _____

Present since _____ *Improving / Unchanging / Worsening*

Commenced as a result of _____ *Or No Apparent Reason*

Symptoms at onset _____ Paraesthesia: *Yes / No*

Spinal history _____ *Cough / Sneeze +ve / -ve*

Constant symptoms: _____ Intermittent Symptoms: _____

Worse *bending sitting / rising / first few steps standing walking stairs squatting / kneeling*
am / as the day progresses / pm when still / on the move Sleeping: *prone / sup / side R / L*
Other _____

Better *bending sitting standing walking stairs squatting / kneeling*
am / as the day progresses / pm when still / on the move Sleeping: *prone / sup / side R / L*
other _____

Continued use makes the pain: *Better Worse No Effect* Disturbed night *Yes / No*

Pain at rest *Yes / No* Site: *Back / Hip / Knee / Ankle / Foot*

Other Questions: *Swelling Clicking / Locking Giving Way / Falling*

Previous episodes _____

Previous treatments _____

General health: *Good / Fair / Poor* _____

Medications: *Nil / NSAIDS / Analg / Steroids / Anticoag / Other* _____

Imaging: *Yes / No* _____

Recent or major surgery: *Yes / No* _____ Night pain: *Yes / No* _____

Accidents: *Yes / No* _____ Unexplained weight loss: *Yes / No*

Summary *Acute / Sub-acute / Chronic* *Trauma / Insidious Onset*

Sites for physical examination *Back / Hip / Knee / Ankle / Foot* *Other:* _____

Version 04/08/2014

EXAMINATION

POSTURE

Sitting *Good / Fair / Poor* Correction of Posture: *Better / Worse / No Effect / NA* Standing: *Good / Fair / Poor*

Other observations: _____

NEUROLOGICAL: **NA / Motor / Sensory / Reflexes / Dural** _____

BASELINES (pain or functional activity): _____

EXTREMITIES ***Hip / Knee / Ankle / Foot***

MOVEMENT LOSS	Maj	Mod	Min	Nil	Pain
Flexion					
Extension					
Dorsi Flexion					
Plantar Flexion					

	Maj	Mod	Min	Nil	Pain
Adduction / Inversion					
Abduction / Eversion					
Internal Rotation					
External Rotation					

Passive Movement (+/- over pressure) (note symptoms and range): _____

PDM	**ERP**

Resisted Test Response (pain) _____

Other Tests _____

SPINE

Movement Loss _____

Effect of repeated movements _____

Effect of static positioning _____

Spine testing *Not relevant / Relevant / Secondary problem* _____

Baseline Symptoms _____

Repeated Tests	**Symptom Response**		**Mechanical Response**	
Active/Passive movement, resisted test, functional test	**During –** Produce, Abolish, Increase, Decrease, NE	**After –** Better, Worse, NB, NW, NE	**Effect –** ↑ or ↓ ROM, strength or key functional test	No Effect
Effect of static positioning				

PROVISIONAL CLASSIFICATION **Extremities** **Spine**

Dysfunction – Articular _____ Contractile _____

Derangement _____ Postural _____

Other _____ Uncertain _____

PRINCIPLE OF MANAGEMENT

Education _____ Equipment Provided _____

Exercise and Dosage _____

Treatment Goals _____

Version 04/08/2014

Appendix E Derangement Vignette

THE McKENZIE INSTITUTE
LOWER EXTREMITIES ASSESSMENT

Vignette ID # : _____2____

Age Range (eg. 35-45 yrs.) : _____**65-75**_____

Gender : M F

Category of Occupation : ____**Admin Assistant**_____

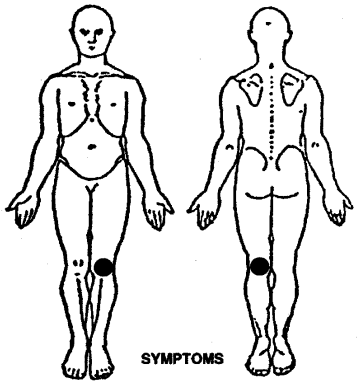
Work : Mechanical stresses _____**up and down, squat, kneeling**__

Leisure : Mechanical stresses _____

Functional disability from present episode_____**Kneeling**__

Functional disability score _____

VAS Score (0-10) _____



HISTORY

Present symptoms _____

Present since **3 Weeks** **Improving** / *Unchanging / Worsening*

Commenced as a result of **Lots of hiking, walking and knee gave way** *Or No Apparent Reason*

Symptoms at onset **Sharp knee pain** Paraesthesia: *Yes* / **No**

Spinal history **Denies** *Cough / Sneeze +ve / -ve*

Constant symptoms: _____ Intermittent Symptoms: **Knee pain**

Worse *bending sitting / rising / first few steps standing walking* **stairs squatting**/*kneeling*
am / *as the day progresses / pm when still / on the move* Sleeping: *prone / sup / side R / L*
Other **Rowing machine, treadmill fast**

Better *bending sitting standing walking stairs squatting / kneeling*
am / as the day progresses / pm **when still** / *on the move* Sleeping: *prone / sup / side R / L*
other **Lying with leg straight**

Continued use makes the pain: *Better* **Worse** *No Effect* Disturbed night **Yes** / *No*

Pain at rest **Yes / No** Site: *Back / Hip / Knee / Ankle / Foot*

Other Questions: *Swelling Clicking / Locking Giving Way / Falling*

Previous episodes **A couple of years ago - resolved**

Previous treatments **None**

General health: **Good** / *Fair / Poor* _____

Medications: *Nil* / **NSAIDS** / *Analg / Steroids / Anticoag / Other* **Ibuprofen**

Imaging: **Yes** / *No* **X-ray – OA changes**

Recent or major surgery: *Yes / No* _____ Night pain: *Yes / No* _____

Accidents: *Yes / No* _____ Unexplained weight loss: *Yes / No*

Summary *Acute* / **Sub-acute** / *Chronic* **Trauma** / *Insidious Onset*

Sites for physical examination *Back / Hip / Knee / Ankle / Foot* *Other:* _____

Version 04/08/2014

EXAMINATION

POSTURE

Sitting *Good / Fair / Poor* Correction of Posture: *Better / Worse / No Effect / NA* Standing: *Good / Fair / Poor*

Other observations: _____

NEUROLOGICAL: **NA / Motor / Sensory / Reflexes / Dural** _____

BASELINES (pain or functional activity): **Pain with squat and steps**

EXTREMITIES ***Hip / Knee / Ankle / Foot***

MOVEMENT LOSS	Maj	Mod	Min	Nil	Pain
Flexion			X		**ERP**
Extension				X	
Dorsi Flexion					
Plantar Flexion					

	Maj	Mod	Min	Nil	Pain
Adduction / Inversion					
Abduction / Eversion					
Internal Rotation					
External Rotation					

Passive Movement (+/- over pressure) (note symptoms and range): _____

	PDM	**ERP**
As active movements		

Resisted Test Response (pain) _____

Not tested

Other Tests _____

SPINE

Movement Loss _____

Effect of repeated movements _____

Effect of static positioning _____

Spine testing **Not relevant** */ Relevant / Secondary problem* _____

Baseline Symptoms _____

Repeated Tests	**Symptom Response**		**Mechanical Response**	
Active/Passive movement, resisted test, functional test	**During –** Produce, Abolish, Increase, Decrease, NE	**After –** Better, Worse, NB, NW, NE	**Effect –** ↑ or↓ ROM, strength or key functional test	No Effect
Rep Extension	Produce ERP	NW	Increase flexion	
Rep OKC Extension	NE	NW		X
Rep Flex with OP	Produce ERP, less with repetition		Increase flexion	
Rep Kneeling	Produce ERP	W		
Effect of static positioning				

PROVISIONAL CLASSIFICATION **Extremities** **Spine**

Dysfunction – Articular _____ Contractile _____

Derangement _____ Postural _____

Other _____ Uncertain _____

PRINCIPLE OF MANAGEMENT

Education _____ Equipment Provided _____

Exercise and Dosage _____

Treatment Goals _____

Version 04/08/2014

Appendix F Articular Dysfunction Vignette

THE McKENZIE INSTITUTE
LOWER EXTREMITIES ASSESSMENT

Vignette ID # : _____1____

Age Range (eg. 35-45 yrs.) : ____**45-55**_____

Gender : M **F**

Category of Occupation : _____

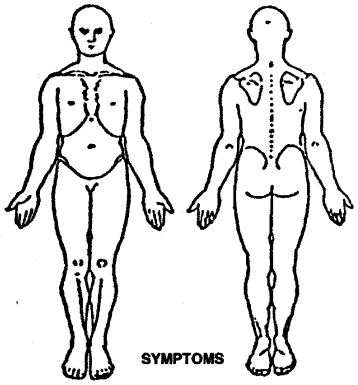
Work : Mechanical stresses ____**Sitting**_____

Leisure : Mechanical stresses _____

Functional disability from present episode_____

Functional disability score _____

VAS Score (0-10) _____



HISTORY

Present symptoms **Knee pain**

Present since **18 months** *Improving / **Unchanging** / Worsening*

Commenced as a result of **Twisted knee** *Or No Apparent Reason*

Symptoms at onset **Knee pain** Paraesthesia: *Yes / No*

Spinal history _____ *Cough / Sneeze +ve / -ve*

Constant symptoms: _____ Intermittent Symptoms: **Knee Pain**

Worse *bending sitting / rising / first few steps standing walking* **stairs** **squatting / kneeling**
am / as the day progresses / pm when still / on the move Sleeping: *prone / sup / side R / L*
Other _____

Better *bending sitting standing walking stairs squatting / kneeling*
am / as the day progresses / pm when still / on the move Sleeping: *prone / sup / side R / L*
other _____

Continued use makes the pain: *Better Worse* **No Effect** Disturbed night *Yes / **No***

Pain at rest *Yes / No* Site: *Back / Hip / Knee / Ankle / Foot*

Other Questions: *Swelling Clicking / Locking Giving Way / Falling*

Previous episodes _____

Previous treatments _____

General health: *Good / **Fair** / Poor* _____

Medications: *Nil / **NSAIDS** / Analg / Steroids / Anticoag / Other* **HTN meds**

Imaging: **Yes** / *No* **X-ray - DJD**

Recent or major surgery: **Yes** / *No* **Menisectomy 16 months ago** Night pain: *Yes / **No***

Accidents: *Yes / **No*** _____ Unexplained weight loss: *Yes / **No***

Summary *Acute / Sub-acute / Chronic* *Trauma / Insidious Onset*

Sites for physical examination *Back / Hip / Knee / Ankle / Foot* *Other:* _____

Version 04/08/2014

EXAMINATION

POSTURE

Sitting *Good / Fair / Poor* Correction of Posture: *Better / Worse / No Effect /* **NA** Standing: *Good / Fair / Poor*

Other observations: _____

NEUROLOGICAL: **NA** / Motor / Sensory / Reflexes / Dural _____

BASELINES (pain or functional activity): **Squat, step down**

EXTREMITIES *Hip* / **Knee** / *Ankle* / *Foot*

MOVEMENT LOSS	Maj	Mod	Min	Nil	Pain
Flexion		X			
Extension				X	
Dorsi Flexion					
Plantar Flexion					

	Maj	Mod	Min	Nil	Pain
Adduction / Inversion					
Abduction / Eversion					
Internal Rotation					
External Rotation					

Passive Movement (+/- over pressure) (note symptoms and range):	PDM	ERP
Extension – full with over pressure		
Flexion – moderate loss		X

Resisted Test Response (pain) **No pain with resisted flexion or extension**

Other Tests _____

SPINE

Movement Loss _____

Effect of repeated movements _____

Effect of static positioning _____

Spine testing **Not relevant** */ Relevant / Secondary problem* _____

Baseline Symptoms _____

Repeated Tests	Symptom Response		Mechanical Response	
Active/Passive movement, resisted test, functional test	**During –** Produce, Abolish, Increase, Decrease, NE	**After –** Better, Worse, NB, NW, NE	**Effect –** ↑ or ↓ ROM, strength or key functional test	No Effect
Rep passive flexion	Produce	NW		X
Rep loaded flexion	Produce	NW		X
Rep active extension	NE	NE		X
Effect of static positioning				

PROVISIONAL CLASSIFICATION **Extremities** **Spine**

Dysfunction – Articular _____ Contractile _____

Derangement _____ Postural _____

Other _____ Uncertain _____

PRINCIPLE OF MANAGEMENT

Education _____ Equipment Provided _____

Exercise and Dosage _____

Treatment Goals _____

Version 04/08/2014

Appendix G Contractile Dysfunction Vignette

THE McKENZIE INSTITUTE
LOWER EXTREMITIES ASSESSMENT

Vignette ID # : 6

Age Range (eg. 35-45 yrs.) : **10-20**

Gender : **M** F

Category of Occupation : **Student**

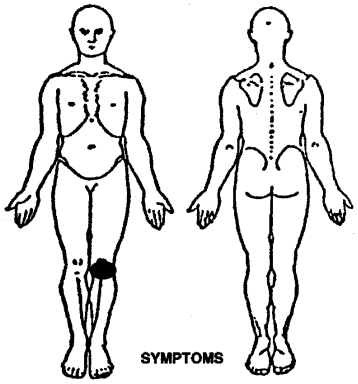
Work : Mechanical stresses **School**

Leisure : Mechanical stresses **Football - Punter**

Functional disability from present episode **Pain with kicking**

Functional disability score

VAS Score (0-10)



HISTORY

Present symptoms **Left anterior knee pain**

Present since **3 weeks** *Improving* / **Unchanging** / *Worsening*

Commenced as a result of **With kicking** *Or No Apparent Reason*

Symptoms at onset **Anterior knee pain** Paraesthesia: *Yes / No*

Spinal history **Denies** *Cough / Sneeze +ve / -ve*

Constant symptoms: Intermittent Symptoms: **Anterior knee pain**

Worse *bending sitting* / **rising** / *first few steps standing walking* **stairs** **squatting** / *kneeling*
am / as the day progresses / pm when still / on the move Sleeping: *prone / sup / side R / L*
Other **Kicking**

Better *bending sitting standing walking stairs squatting / kneeling*
am / as the day progresses / pm when still / on the move Sleeping: *prone / sup / side R / L*
other

Continued use makes the pain: *Better* **Worse** *No Effect* Disturbed night *Yes* / **No**

Pain at rest **Yes** – *after kicking / No* Site: *Back / Hip* / **Knee** / *Ankle / Foot*

Other Questions: *Swelling Clicking / Locking Giving Way / Falling*

Previous episodes **Last football season**

Previous treatments **None**

General health: **Good** / *Fair / Poor*

Medications: **Nil** / *NSAIDS / Analg / Steroids / Anticoag / Other*

Imaging: *Yes* / **No**

Recent or major surgery: *Yes* / **No** Night pain: *Yes* / **No**

Accidents: *Yes* / **No** Unexplained weight loss: *Yes* / **No**

Summary *Acute* / **Sub-acute** / *Chronic* **Trauma** / *Insidious Onset*

Sites for physical examination *Back / Hip* / **Knee** / *Ankle / Foot* *Other.*

Version 04/08/2014

EXAMINATION

POSTURE

Sitting *Good / Fair / Poor* Correction of Posture: *Better / Worse / No Effect / NA* Standing: *Good / Fair / Poor*

Other observations: _____

NEUROLOGICAL: **NA / Motor / Sensory / Reflexes / Dural** _____

BASELINES (pain or functional activity): **Squatting** _____

EXTREMITIES ***Hip / Knee / Ankle / Foot***

MOVEMENT LOSS	Maj	Mod	Min	Nil	Pain
Flexion				X	X
Extension				X	
Dorsi Flexion					
Plantar Flexion					

	Maj	Mod	Min	Nil	Pain
Adduction / Inversion					
Abduction / Eversion					
Internal Rotation					
External Rotation					

Passive Movement (+/- over pressure) (note symptoms and range):	**PDM**	**ERP**
No loss of extension with over pressure		
No loss of flexion with over pressure		X

Resisted Test Response (pain) **Pain with resisted extension**
Pain free resisted flexion

Other Tests _____

SPINE

Movement Loss _____

Effect of repeated movements _____

Effect of static positioning _____

Spine testing *Not relevant / Relevant / Secondary problem* _____

Baseline Symptoms **None**

Repeated Tests	**Symptom Response**		**Mechanical Response**	
Active/Passive movement, resisted test, functional test	**During –** Produce, Abolish, Increase, Decrease, NE	**After –** Better, Worse, NB, NW, NE	**Effect –** ↑ or ↓ ROM, strength or key functional test	No Effect
Rep. Passive Extension	NE	NE		NE
Rep. Active Extension	Produce	NW		NE
Rep. Flexion with OP	Produce	NW		NE
Effect of static positioning				

PROVISIONAL CLASSIFICATION **Extremities** **Spine**

Dysfunction – Articular _____ Contractile _____

Derangement _____ Postural _____

Other _____ Uncertain _____

PRINCIPLE OF MANAGEMENT

Education _____ Equipment Provided _____

Exercise and Dosage _____

Treatment Goals _____

Version 04/08/2014

Appendix H OTHER Vignette

THE McKENZIE INSTITUTE
LOWER EXTREMITIES ASSESSMENT

Vignette ID # : ____10_____

Age Range (eg. 35-45 yrs.) : _____**15-25**_____

Gender : **M** F

Category of Occupation : _____**Student**_____

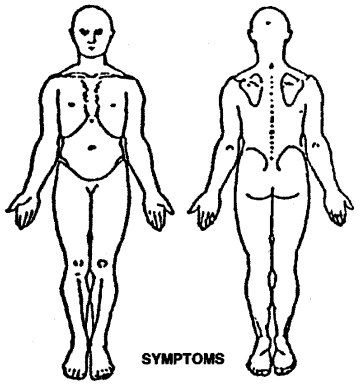
Work : Mechanical stresses _____

Leisure : Mechanical stresses _____**Soccer**_____

Functional disability from present episode_____

Functional disability score _____

VAS Score (0-10) _____



HISTORY

Present symptoms **Knee pain**

Present since **2 weeks** **Improving** / *Unchanging / Worsening*

Commenced as a result of **Twist playing soccer** *Or No Apparent Reason*

Symptoms at onset _____ Paraesthesia: *Yes / No*

Spinal history _____ *Cough / Sneeze +ve / -ve*

Constant symptoms: _____ Intermittent Symptoms: **Knee Pain**

Worse *bending sitting / rising / **first few steps** standing walking **stairs** **squatting / kneeling***
am / as the day progresses / pm when still / on the move Sleeping: *prone / sup / side R / L*
Other _____

Better *bending sitting standing walking stairs squatting / kneeling*
*am / as the day progresses / pm when still / **on the move*** Sleeping: *prone / sup / side R / L*
other _____

Continued use makes the pain: **Better** *Worse No Effect* Disturbed night *Yes / **No***

Pain at rest *Yes / **No*** Site: *Back / Hip / Knee / Ankle / Foot*

Other Questions: **Swelling** *Clicking / Locking* *Giving Way / Falling*

Previous episodes _____

Previous treatments **Ice**

General health: **Good** / *Fair / Poor* _____

Medications: *Nil / **NSAIDS** / Analg / Steroids / Anticoag / Other* _____

Imaging: *Yes / **No*** _____

Recent or major surgery: *Yes / **No*** _____ Night pain: *Yes / **No*** _____

Accidents: *Yes / **No*** _____ Unexplained weight loss: *Yes / No*

Summary *Acute / Sub-acute / Chronic* *Trauma / Insidious Onset*

Sites for physical examination *Back / Hip / **Knee** / Ankle / Foot* *Other:* _____

Version 04/08/2014

EXAMINATION

POSTURE

Sitting *Good / Fair / Poor* Correction of Posture: *Better / Worse / No Effect / NA* Standing: *Good / Fair / Poor*

Other observations: _____

NEUROLOGICAL: ***NA*** */ Motor / Sensory / Reflexes / Dural* _____

BASELINES (pain or functional activity): **Squat, steps** _____

EXTREMITIES Hip / **Knee** / Ankle / Foot

MOVEMENT LOSS	Maj	Mod	Min	Nil	Pain
Flexion			**X**		**ERP**
Extension			**X**		**ERP**
Dorsi Flexion					
Plantar Flexion					

	Maj	Mod	Min	Nil	Pain
Adduction / Inversion					
Abduction / Eversion					
Internal Rotation					
External Rotation					

Passive Movement (+/- over pressure) (note symptoms and range): _____

	PDM	**ERP**
Flexion and extension WNL		X

Resisted Test Response (pain) **Quads and hams 4+/5** _____

Other Tests _____

SPINE

Movement Loss _____

Effect of repeated movements _____

Effect of static positioning _____

Spine testing **Not relevant** */ Relevant / Secondary problem* _____

Baseline Symptoms _____

Repeated Tests	**Symptom Response**		**Mechanical Response**	
Active/Passive movement, resisted test, functional test	**During –** Produce, Abolish, Increase, Decrease, NE	**After –** Better, Worse, NB, NW, NE	**Effect –** ↑ or↓ ROM, strength or key functional test	No Effect
Rep flexion	Produce	NW		X
Rep extension	Produce	NW		X
Effect of static positioning				

PROVISIONAL CLASSIFICATION **Extremities** **Spine**

Dysfunction – Articular _____ Contractile _____

Derangement _____ Postural _____

Other _____ Uncertain _____

PRINCIPLE OF MANAGEMENT

Education _____ Equipment Provided _____

Exercise and Dosage _____

Treatment Goals _____

Version 04/08/2014

Appendix I Phase 2 Recruitment Email

Dear MDT Clinician,

You are invited to participate in a research study that is investigating the inter-examiner reliability of the McKenzie System of MDT when used by MDT trained clinicians in patients presenting with musculoskeletal knee pain. For this study, 6 expert MDT raters are required to review and classify 53 clinical vignettes.

We are asking you because you are Credentialed and/or have a Diploma standing within the McKenzie Institute and have been applying MDT to musculoskeletal problems in the extremity for greater than 3 years.

This study will be conducted by Trevor Birmingham, a Professor in the School of Physical Therapy at Western University in the School of Physical Therapy. Sean Willis, a Master of Science student in Health and Rehabilitation Sciences at Western University, Shawn Robbins, an Assistant Professor in the School of Physical and Occupational Therapy at McGill University and Richard Rosedale, an Instructor with the McKenzie Institute will also be participating in the study.

Attached is a Letter of Information for you to review and consider your participation. Also attached is a Consent Form for you to sign and submit should you wish to participate in this study.

Thank you for taking the time to consider participation in this study.

Sincerely,

Sean Willis

Appendix J Phase 2 Letter of Information and Consent Form

School of Physical Therapy

Letter of Information and Consent Form

Classification of Clinical Vignettes

Inter-observer Reliability of the McKenzie System of Mechanical Diagnosis and Therapy in the Examination of the Knee

Study Principal Investigator
Trevor Birmingham PhD, BSc(PT)
Western University, School of Physical Therapy

Study Co-Investigators
Sean Willis BSc(PT), MSc Student
Western University, Health and Rehabilitation Sciences

Shawn Robbins PhD, BSc(PT)
McGill University, School of Physical and Occupational Therapy

Study Advisor
Richard Rosedale PT, Dip MDT
Instructor, McKenzie Institute

1. Introduction

You are invited to participate in a research study that is investigating the inter-examiner reliability of the McKenzie System of MDT when used by MDT trained clinicians in patients presenting with musculoskeletal knee pain. For this study, 6 expert MDT raters are required to

review and classify 53 clinical vignettes. The purpose of this letter is to provide you with the necessary background and information to make an informed decision regarding your participation within this study.

2. Why Are We Asking You?

We are asking you because you are Credentialed and/or have a Diploma standing within the McKenzie Institute and have been applying MDT to musculoskeletal problems in the extremity for greater than 3 years.

We are giving this letter of information only to people who are registered on the publically available list of MDT clinicians practicing in Canada and the United States, who understand both written and spoken English and who did not participate in generating the 53 clinical vignettes in the study. If the above situation does not apply to you, we ask that you do not volunteer to participate in our study.

3. Who is Conducting the Study?

This study will be conducted by Trevor Birmingham, a Professor in the School of Physical Therapy at Western University in the School of Physical Therapy. Sean Willis, a Master of Science student in Health and Rehabilitation Sciences at Western University, Shawn Robbins, an Assistant Professor in the School of Physical and Occupational Therapy at McGill University and

Richard Rosedale, an Instructor with the McKenzie Institute will also be participating in the study.

4. What will happen if you agree?

If you agree to participate, Sean Willis will contact you via email with 53 clinical vignettes and instructions regarding classifying each of the vignettes into one of four mechanical syndromes: derangement, articular dysfunction, contractile dysfunction and other.

In addition to completing the classification of the clinical vignettes, you will also be asked to complete and submit a data collection form to Sean Willis that will record your: gender, age, country of practice, clinical practice setting, years of practice, length of time Credentialed/Diploma, discipline (eg. PT versus DC), proportion of extremity patients treated and proportion of knee patients treated with MDT. This information will be used within the study to describe the clinical background and characteristics of the group of expert raters as well as analyze the effect those characteristics may have on the level of agreement between raters.

It is expected that the review and subsequent rating of the 53 clinical vignettes may take some time. The anticipated time commitment to review and rate each clinical vignette is 30 minutes for a total of 27 hours to rate all vignettes. A total of 2 months will be provided for this process.

5. What are the risks and benefits?

There are no direct benefits from participating in this study. However, for clinicians who treat people with musculoskeletal knee disorders, this would be the first study to address inter-examiner reliability of McKenzie system in assessing patients with knee pain. It would provide MDT practitioners with valuable information about reliability of this assessment method.

There are no risks to participating in this study.

6. What about confidentiality?

Your confidentiality will be protected. No information that discloses your identity will be released or be a part of any publication of the results of the study. Your first and last name as well as your address, email address and phone number will be collected for contact and communication purposes only. All records will be given a code ID to be used on all collected forms. All information collected will be kept in locked filing cabinets and shredded seven years after the completion of the study.

7. Alternatives to participation?

Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions or withdraw from the study at any time with no effect on your future employment.

Should you withdraw from the study, you may also withdraw any information or data that you have submitted.

8. If you agree to participate in the study

Please sign the attached consent form and return to Sean Willis by email or fax. You do not waive any of your legal rights by signing the consent form.

You may keep this letter of information. A copy of your signed consent form will be made for you. If you have any questions about this study, please contact Sean Willis.

If you have any questions about your rights as a research participant or the conduct of the study, you may contact The Office of Research Ethics at Western University.

Representatives of The University of Western Ontario Health Sciences Research Ethics Board may contact you or require access to your study-related records to monitor the conduct of the research.

School of Physical Therapy

Consent Form

Inter-observer Reliability of the McKenzie System of Mechanical Diagnosis and Therapy in the Examination of the Knee

I have read the accompanying letter of information and I agree to participate. Questions I had have been answered to my satisfaction.

Date	Participant's Name	Participant's Signature
_____	_____	_____

Date	Name of Person Obtaining Informed Consent	Signature of Person Obtaining Informed Consent
_____	_____	_____

Initials_____

LOI 08/08/2014

Appendix K Phase 2 Data Collection Form

Subject Characteristics

Age:_____ Years of Practice:_____ Diploma/Credentialed

Clinical Practice Setting: Private Practice

Insurance/Workman's Compensation

Hospital Outpatient Setting

Physician Referral

Discipline: Physical Therapist

Chiropractor

Proportion of Caseload that are Extremity Patients:_____

Proportion of Knee patients treated with MDT:_____

Contact Information

Name (Please Print):

Address:

City:

Email:

Phone Number (H): (W):

Curriculum Vitae

Name:	Sean Willis
Post-secondary Education and Degrees:	The University of Western Ontario London, Ontario, Canada 1994 - 1998 BSc (PT)
Post-graduate Qualifications:	MDT Credentialing Exam April 2008
Related Work Experience	Physiotherapist London Health Sciences Centre 2000 - Present Clinical Associate Western University, School of Physical Therapy 2005 - Present Teaching Assistant Western University 2015